Annotated Bibliography of Evaluating the Educational Impact of Digital Libraries

-- Compiled by Sarah Giersch, Kirsten Butcher & Tom Reeves

National Science Digital Library

September 2003

NSF Grant No. 0127298

This material is based upon work supported by the National Science Foundation under Grant No. 0127298. 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, the University Corporation for Atmospheric Research or The University of Georgia.

Table of Contents

Introduction

This annotated bibliography was commissioned to support the NSDL Evaluation Workshop (planned for October 2003) that will 1) explore the issues around evaluating the impact of digital libraries on education and that will 2) begin developing a strategy to evaluate the impact of the National Science Digital Library (NSDL) on STEM education.

The bibliography's purpose, then, is to identify research to date on evaluating the impact of digital libraries on learning and teaching. It contains a *Summary & Analysis* section, which defines terms, makes observations about the literature, reviews the resources included, highlights issues, and suggests areas for further consideration. The A*nnotated Bibliography* is organized alphabetically. Details about each resource include: a citation; the abstract provided with each resource; a summary, if necessary, to highlight relevant ideas not expressed in the abstract; and, a comment section with a short analysis of the resource. The annotations are followed by a description of the M*ethodology* used to compile this document.

Summary & Analysis

Terms to Define

Digital library: Definitions of this term vary among communities, primarily computer science and information and library science, who pioneered the development and application of digital library technologies for different purposes. The debate to define digital libraries will continue as the population of digital library end-users expands beyond the originating community. Over the past four years, Borgman (2002) [1] has examined the proliferating definitions of the terms digital library and has developed a two-part definition:

1. Digital libraries are a set of electronic resources and associated technical capabilities for creating, searching, and using information. In this sense, they are an extension and enhancement of information storage retrieval systems that manipulate digital data in any medium (text, image, sound; static or dynamic images) and exist in distributed networks. The content of digital libraries includes data, metadata that describe various aspects of the data (e.g., representation, creator, owner, reproduction rights), and metadata that consist of links or relationships to other data or metadata, whether internal or external to the digital library.
2. Digital libraries are constructed – collected and organized – by [and for] a community of users, and their functional capabilities support the information needs and uses of that community. They are a component of communities in which individuals and groups interact with each other, using data, information, and knowledge resources and systems. In this sense they are an extension, enhancement, and integration of a variety of information institutions as physical places where resources are selected, collected, organized, preserved, and accessed in support of a user community. These

[1] Borgman, C. (2002). Challenges in building digital libraries for the 21st Century. Digital Libraries: People, Knowledge & Technology: Proceedings of the 5th International Conference on Asian Digital Libraries (ICADL 2002), Singapore, December 12-14, 2002, Heidelberg: Springer-Verlag. Online: http://www.springer.de/cgi/svcat/search_book.pl?isbn=3-540-00261-8

> information institutions include, among others, libraries, museums, archives, and schools; but digital libraries also extend and server other community settings, including classrooms, offices, laboratories, homes and public spaces. (Borgman, 2002, p. 2)

Technology: Early in the development of this bibliography, the authors expanded the disciplines reviewed to include education. The purpose was not to find studies on the impact of all learning technologies, but instead to incorporate hypotheses, methodologies, and perhaps baseline data that would assist in framing the evaluation of digital libraries. So, rather than debate the ultimate definition of "technology," preceded by the appropriate adjective (education, learning etc.), keep in mind Reeves' (1998) definition, which acknowledges the many different definitions of technology, and media, and then distinguishes between them as follows: "With respect to education, *media* are the symbol systems that teachers and students use to represent knowledge; *technologies* are the tools that allow them to share their knowledge representations with others" (p. 1).

Impact: While *defining* impact is relatively easy (*n.* 3. The effect or impression of one thing on another[2]), the literature from all the disciplines reviewed indicates that researchers are still grappling with how to structure evaluation studies to *identify* impact. Oliver and Harvey (2002) observe that "…most projects concerned with the introduction of new technology… aim to have some kind of impact on students. Often this is framed in terms of an impact on learning, although perhaps more commonly observed are changes in behaviour" (p. 19). However, studies in this bibliography also evaluated the impact of technology on other groups such as teachers, administrators, and more broadly, organizations and infrastructure, while expecting impact to manifest in a variety of behaviors, attitudes and skills. Contextual features often complicate identifying the impact of technology.

Observations about the Literature

To identify resources for this annotated bibliography, literature from the following fields was consulted: cognitive psychology, computer science, education, education policy, information and library science. The document includes resources from each discipline, in part, because each field offers concepts and methodologies that should be considered when evaluating the impact of digital libraries on teaching and learning, and in part, because no one field contains a critical mass of resources on this topic. While this might be perceived as a gap in the literature, the lack of resources could also be attributed to the particular focus of a discipline. For example, library science isn't necessarily interested in the long-term cognitive effects of traditional library services on users. Similarly, digital libraries are a relatively new technology and have not been widely adopted by teachers, hence there has been no research to evaluate its impact on teaching and learning. Instead of identifying gaps, then, the following points briefly detail the prevalence of research on digital libraries, learning technologies and studies of impact in the disciplines reviewed.

- *Cognitive psychology* literature is lacking in quality reviews of educational technology and its effectiveness. Articles dealing with educational technology tend to focus on a specific type of tool/intervention and a particular set of outcome analyses. This focus on (short-term) outcomes is particularly

[2] The American Heritage® Dictionary of the English Language, Fourth Edition Copyright © 2000 by Houghton Mifflin Company. Retrieved from www.dictionary.com 9/14/03.

problematic to identifying effects of educational technology on learning behaviors. There is little attention to how students use educational technology and how such use relates to changes in learning processes as well as to learning outcomes. Further, studies (likely for practical reasons) usually do not attempt to track long-term changes in technology use, learning processes, or outcomes associated with educational technology.

- *Computer science* literature is less interested in the educational impact of digital libraries, and is focused instead on evaluating the usability of the technology.
- Reviews of *education* literature identified studies on the impact of learning technologies, with more emphasis placed on how learning technologies affect teaching and to a lesser researched extent, learning. Also, in the education literature, digital libraries are just one manifestation of learning technologies.
- *Information and library science* literature produced a few studies on the impact of traditional libraries on learning, though most of the literature is devoted to evaluating the impact of technology provided by traditional libraries (e.g., electronic databases) on the search and retrieval process.
- Literature on digital libraries is often found in the disciplines of its developer communities, namely computer science and information and library science. However, the number of journals dedicated to reporting digital library research is continually expanding. Most studies reporting evaluative work focus on a design and development to improve technical systems or digital library features and services.

Literature Review & Findings

There is a growing body of literature, primarily from the fields of computer science and information and library science, about evaluating digital libraries. As digital libraries have evolved, so have the documented efforts of evaluating system design, implementation and use. For the most part, early research examined *if* and *how* people were using digital libraries; this data was fed into the design and development cycle. Very little work has been done to evaluate the effects of using digital libraries in the context of teaching, learning or research. Barriers to digital library evaluation efforts include the relative new-ness of the field and the rapid pace of development, factors which preclude extensive summative evaluation. Further hindering efforts is a lack of tools and methodologies such that "the conceptual state-of-the-art of digital library evaluation is not sufficiently developed to start with" (Saracevic, 2000, p. 352). I would add that limited involvement of people with evaluation expertise and limited funding have allowed few digital library development efforts to take a longitudinal and comprehensive approach toward the type of evaluation that addresses the impact and effects digital libraries have on their users. Two exceptions are the ADEPT Digital Library (Borgman, 2000) and the Perseus Digital Library (Marchionini, 2000). From the very beginning, both incorporated formative and summative evaluations into every aspect of digital library development, made studying the impact on end-users' learning a priority and had funding for a duration that provided an opportunity to comprehensively study the effects of digital libraries.

Although they use different metrics and methods of data gathering, traditional libraries evaluate features that are paralleled in digital libraries, such as library use, effectiveness of services and patron demographics. Despite a long history and a solid body of literature about traditional library evaluation efforts, only recently have traditional, primarily academic, libraries begun to consider how to incorporate new

measures that might be "surrogates of impact" (Everest, 2001, p. 19). At the same time, Everest (2001) argues that evaluation "will be made even more complicated as teaching, learning and research become even more integrated with libraries' resources and services" (p. 22).

The dearth of research on specific studies evaluating the impact of digital or traditional libraries on learning, for reasons noted above, warranted expanding the disciplines reviewed to encompass research done in the field of education, on the assumption that digital libraries are a type of learning technology and could take advantage of the methodologies used in evaluation research on learning technologies.

Three comprehensive reports (Sivin-Kinchala, 2000; Murphy, et al., 2002; Reeves 1998) examined hundreds of studies from the past decade that sought to identify the effectiveness and impact of a variety of hardware and software applications in a variety of learning environments (k-12, higher education, adult education). While all three reports concluded that technology does have a positive effect on learning, Reeves noted that "pedagogy… is most influential on learning, not media or technology" (p. 3). In compiling their report, Murphy et al., (2002) included only research that met standards of methodological quality and relevance, which limited the number of reports used and led researchers to conclude that,

> "without a concerted effort by the field to improve the quality of evaluation research, the current evidence-based knowledge will remain insufficient to guide policy-makers, decision-makers, practitioners, and designers in the field" (p. 3).

Despite this recommendation, observations made while preparing this bibliography suggest that the quality of research on learning technology has improved, but very slowly. As a result, only two specific studies were included, primarily as examples of two different approaches in evaluating the effectiveness and impact of learning technology: experimental design (Tzortzidou, 2001) and case study (Gunn, 2002).

Even though research to evaluate the impact of technology on learning has not improved, teachers, researchers and policy-makers continue to design, implement and fund studies for a multitude of reasons. However, there is a middle ground of literature, between the micro, situation- and technology-specific studies and the macro meta-analyses of multiple studies, which contains reports and plans of evaluation efforts on a program or system level. These documents function as potential models, or at a minimum checklists, for developing strategic plans. They provide definitions, identify stakeholders and document processes, which offer a perspective that is not expressed or implemented in the micro and macro studies.

Although developed primarily in the context of learning technology, the issues identified and described below are just as applicable to consider when developing a plan to evaluate the impact of digital libraries on learning.

- Adopting a long-term approach to gathering data as part of an evaluation plan is critical to identifying meaningful results. Long-term studies, Marchionini's (2000) in particular, noted that short-term, spot evaluation results were not as relevant or meaningful as long-term evaluation results.
- Teachers are a crucial part of the context that affects student learning. However, measures of impact should distinguish between technology's effects on teachers

and on students. Several studies reviewed for this bibliography measured the impact of technology on teachers; however, some studies used this measure to infer that student learning had occurred.
- Current standardized assessments used to ascertain whether learning has occurred do not reflect the role of technology in the teaching and learning process. This results in the development of "localized" assessments geared towards a particular application of technology.
- Evaluation studies that produced meaningful (i.e., a definitive answer to research questions, though not necessarily generalizable) results incorporated quantitative and qualitative measures of data.
- In order for results to be generalized, evaluation design and reporting should adhere to accepted standards, which could be challenging given the inter-disciplinary nature of this topic.

Items for Further Consideration

The focus of this annotated bibliography precludes further research on an interesting and convoluted topic; however, there is much more work that could be done. For example, classifying all of the literature examined, but not necessarily included, in this bibliography by categories such as technology (or types of digital library), context of use, approaches to learning and users would help researchers in all fields towards a common understanding of the issues and avenues for further work. As noted above, the particular topic of evaluating the impact of digital libraries on learning is not within the purview of the particular disciplines mentioned and, therefore, has received incidental attention. As the fields of computer science and information and library science pioneered the development of digital libraries, they could also lead the way, with insights from education and cognitive psychology, in designing evaluations about the impact of that technology. Further research could only benefit from a continued multi-disciplinary approach.

Annotated Bibliography

Adams, A. and A. Blandford (2003). "The Unseen and Unacceptable Face of Digital Libraries." *Journal of Digital Libraries*(forthcoming). Online: http://www.uclic.ucl.ac.uk/annb/DLUsability/AAABjdl.pdf

Author Abstract: The social and organizational aspects of digital libraries are often overlooked but this paper reviews how they can affect users' awareness and acceptance of digital libraries. An analysis of research conducted within two contrasting domains (Clinical and Academic) is presented which highlights issues of user interactions, work practices and the organizational social structures. The combined study comprises an analysis of 98 in-depth interviews and focus groups with lecturers, librarians and hospital clinicians. The importance of current and past roles of the library, and how users interacted with it, are revealed. Web-based digital libraries, while alleviating most library resource and interaction problems, require a change in librarians' and DL designers' roles and interaction patterns if they are to be implemented acceptably and effectively. Without this role change, users will at best be unaware of these digital resources and at worst feel threatened by them. The findings of this paper highlight the importance on DL design and implementation of the social context and supporting user communication (i.e. collaboration and consultation) in their information search and usage activities. [39 references]

Summary: Comprehensive abstract; no summary needed.

Comment: Digital library use has grown beyond the communities that initially developed them. As with any technology, early development efforts produced, in this case, digital libraries tailored to community-specific needs or early adopter requirements. The social issues that Adams and Blandford identify (knowledge-hoarding, reliance on traditional information dissemination methods and changing roles and expectations) could be applicable to many "second generation" communities adopting new technology. As noted by this and other articles in the bibliography, the impact of technology (digital libraries, course management systems or spreadsheets) cannot be considered independently of the various contexts, including social and organizational, in which they are implemented.

Agodini, R., M. Dynarski, et al. (2003). The Effectiveness of Educational Technology: Issues and Recommendations for the National Study - Draft. Princeton, NJ, Mathematica Policy Research, Inc.**:** 51 pages. Online: http://www.ed.gov/about/offices/list/os/technology/issues.pdf

No Abstract. [6 references]
Summary: This draft report describes a research plan to answer the question "Is educational technology effective in improving student academic achievement?" as mandated by the No Child Left Behind Act. Based on issues and strategies identified by an expert review panel, three institutions (Mathematica Policy Research, Inc. and its partners the American Institute for Research and the Education Development Center) working with the U.S. Department of Education (ED) developed this report which formulated recommendations; considered how to address factors that impact technology effectiveness such as teacher technology training and student learning context (i.e., parents, neighborhoods and student characteristics); and, described

approaches to designing and conducting the study including how to select technology applications to study and schools to participate.
The expert review panel posed four questions, which the research partners answered in the first part of the report in the form of recommendations. These serve to focus the direction and scale of the study.
Question 1: *What is "educational technology?": Recommendation 1: Examine technology applications (rather than hardware) designed to support teaching and learning. Recommendation 2: Use a public submission process to identify technology applications to study.* The design team further decided to group applications by type and to classify potential applications based on skills targeted, instructional approach supported, intensity in terms of frequency and duration of use, and amount of professional development required.
Question 2: *What is "effective?": Recommendation 3: Use experimental designs to measure effects. Recommendation 4: Study the effects of technology applications for schools or teachers that do not currently use the applications but are interested in using them. Recommendation 5: Design the study to detect "moderate" to "large" effects of technology applications.* These are explored in more detail in the second part of the report.
Question 3: *What kinds of students?: Recommendation 6: Study the effects of technology applications for students in the primary and secondary grade levels (K-12). Recommendation 7: Study the effects of technology applications for schools that receive Title I funds.*
Question 4: *What is "academic achievement?": Recommendation 8: Study the effects of technology applications on student academic achievement as measured by commonly used standardized tests, and collect data on other academic indicators to provide a fuller picture. Recommendation 9: Study the effects of technology applications that support instruction in reading and math.*

The second part of the report considers approaches to studying the effectiveness of technology in response to Question 2 and its recommendations. The design team developed a general conceptual framework showing the links between technology applications and learning. Next the authors explored issues related to random sampling. The design team decided to use a control and treatment group, but weighs (and doesn't reach a conclusion about) which unit to sample (student, teacher, classroom, school), recognizing that this choice will be driven by the technology applications implemented and schools' ability to support randomly assigning students to classes. The authors further consider the best unit to sample in terms of a power analysis; and, they discuss issues about estimating the impacts of technology applications.

Comment: This report presents an approach to education research supported by the current federal administration. How could this be applied to digital libraries?

Baylor, A. and D. Ritchie (2002). "What factors facilitate teacher skill, teacher morale, and perceived student learning in technology-using classrooms?" *Computers & Education* **39**(4): 395-414.

Author Abstract: Based on a comprehensive study of 94 classrooms from four states in different geographic regions of the country, this quantitative study investigated the impact of seven factors related to school technology (planning, leadership, curriculum alignment, professional development, technology use, teacher openness

to change, and teacher non-school computer use) on five dependent measures in the areas of teacher skill (technology competency and technology integration), teacher morale, and perceived student learning (impact on student content acquisition and higher order thinking skills acquisition). Stepwise regression resulted in models to explain each of the five dependent measures. Teacher technology competency was predicted by teacher openness to change. Technology integration was predicted by teacher openness to change and the percentage of technology use with others. Teacher morale was predicted by professional development and constructivist use of technology. Technology impact on content acquisition was predicted by the strength of leadership, teacher openness to change, and negatively influenced by teacher non-school computer use. Technology impact on higher-order thinking skills was predicted by teacher openness to change, the constructivist use of technology, and negatively influenced by percentage of technology use where students work alone. Implications for the adoption and use of school technologies are discussed. [27 references]

Summary: This paper provides a highly detailed description of the study's methodology, which took place in k-12 classrooms across the US. The framing question for the study was, "What actions can school personnel take that most effectively lead to their desired results regarding the integration of technology in schools?" The authors report how independent and dependent variables were developed from work in educational research, cognitive psychology and educational technology. Soliciting teacher and administrator input was key to operationalizing variable definitions. The authors gathered data on each variable through a combination of questionnaires, structured interviews and document review. Data was quantified and analyzed using stepwise regression analysis.

Comment: Unlike other papers in this bibliography, this study did not evaluate a specific hardware or software application. Nor was the context for the evaluation limited to one or two classrooms. This broader scope and scale could serve as a possible model for selecting the various contexts in which digital libraries are used and evaluated. Similarly, the variables identified for this study should be reviewed, along with the process of involving stakeholders (teachers, administrators) in operationalizing definitions. Also, this is one of the few studies that specifically strives to measure technology impact on higher order thinking skills. The drawbacks of this study lie in the fact that 1) data was gathered primarily from teachers, administrators and school documents; no data was gathered specifically from students. This reflects, in part, the focus of the study; and, 2) though the authors used a variety of instruments to collect data, the results were eventually quantified; this resulted in rapid analysis, but the authors could only make suppositions in their interpretations without the benefit of supporting qualitative results.

Borgman, C., A. Gilliland-Swetland, et al. (2000). "Evaluating digital libraries for teaching and learning in undergraduate education: A case study of the Alexandria Digital Earth ProtoType (ADEPT)." *Library Trends* **49**(2): 228-250. Online: http://is.gseis.ucla.edu/adept/pubs/lt822.html

Author Abstract: This is a discussion on the research design for an educational evaluation of the Alexandria Digital Earth ProtoType (ADEPT), a digital library of gee-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. [31 references]

Summary: This paper describes the research design for the evaluation of ADEPT, a digital library of geo-referenced information resources. The premise is that "digital library services will contribute positively to undergraduate instruction and to student learning of scientific processes." The plan calls for formative evaluations throughout the project to develop design requirements and summative evaluation to assess learning outcomes. Since the project was in its first year when this paper was published, most work on evaluation was directed towards developing prototypes of the I-scapes (Information Landscapes) tool for expressing and visualizing geo-spatial concepts and processes. Assessing learning outcomes will occur later in the project, though the paper presented the methodology and design issues involved in the summative evaluation plan. The researchers wanted to know "the cognitive consequences of participating in an ADEPT environment." They plan to test a group using ADEPT and a comparison group on performance tasks requiring skills specific to geographical thinking. The researchers had already given some thought to addressing the issues of reliability, validity, standardization and objectivity when designing the performance assessments. Specifically, they included scripted guidance for testing over problems which were related to those studied in class using several geography scenarios (instead of one) and using quantitative measures. The rest of the paper describes first-year progress consisting of results of classroom observation and deployment of the I-scape prototype.

Comment: Very few digital library projects have developed plans that make formative and summative evaluation such an integral part of their research. This can serve as a model for other projects developing digital libraries, even though most projects have a span of 2-3 years rather than five. One cautionary note, though, is that tailoring performance assessments to a specific piece of technology could render the results (though perhaps not the development process) irrelevant as educational technology, digital libraries rapidly evolve, accompanied by an evolution in teacher and learner knowledge and skills.

Chinn, C. A. and W. F. Brewer (2001). "Models of data: A theory of how people evaluate data." *Cognition and Instruction* **19**(3): 323-393.

Author Abstract: This article reports the results of a study investigating how undergraduates evaluate realistic scientific data in the domains of geology and paleontology. The results are used to test several predictions of a theory of data evaluation, which we call models-of-data theory. Models-of-data theory assumes that when evaluating data, the individual constructs a particular kind of cognitive model that integrates many features of the data with a theoretical interpretation of the data. The individual evaluates the model by attempting to generate alternative causal explanations for the events in the model. We contrast models-of-data theory with other proposals for how data are cognitively represented and show that models-of-data theory gives a good account of the pattern of written evaluations of data

produced by the undergraduates in the study. We discuss theoretical and instructional implications of the theory.

Summary: This article provides a good theoretical basis for evaluating the ways in which students work with data. Chinn and Brewer use a relatively simple experimental design in which students:

1. Read about the initial theory
2. Read about the data
3. Rate their belief in the data and explain their ratings.
4. Rate the consistency of data with theory and explain these ratings.

Results are generally consistent with the models-of-data theory and suggest implications for instruction and assessment when students are learning with scientific data. Specifically, Chinn and Brewer argue that their results suggest that a series of strategies could improve students' evaluation of scientific data. These strategies are (p. 381):

1. Constructing an accurate model (of the presented data).
2. Elaborating the model (to achieve a complete representation of the presented data).
3. Elaborating the model to include additional information.
4. Recognizing when one's model is insufficiently elaborated.
5. Searching for alternative causes.
6. Using covariation evidence.
7. Using strategies for evaluating inductive and analogical generalizations.
8. Refraining from accepting implausible alternative causes.
9. Avoiding the use of nonspecific alternative causes.
10. Avoiding biased denial of data.
11. Avoiding biased acceptance of data.

Chinn and Brewer suggest that teaching students to explicitly map out their models of scientific data could be used as an instructional method but also acknowledge that learning to evaluate scientific evidence also requires that the student learn a great deal about the domain from which the data comes.

Comment: The methodologies and theory described in this paper suggest a number of assessment strategies for evaluating the educational impact of authentic data that is included in digital libraries. It also addresses a central concern of science instruction: understanding and evaluating the ways in which students understand and work with scientific data.

Everest, K. and P. Payne (2001). "The impact of libraries on learning, teaching and research: Report on the LIRG seminar held in Leeds, 30th. October 2001." *Library and Information Research News* **25**(81): 18-22.

Author Abstract: Report of a one day seminar held in Leeds, UK, by the Library and Information Research Group in October 2001, to consider the measuring of the impact of academic libraries on learning, teaching and research. Summarizes presentations by Peter Brophy, on methods, approaches and outcomes of measuring the impact of libraries and David Streatfield, on experience with projects that have used measures of impact or effectiveness. Also describes three university library case studies and presents the seminar's conclusions. [1 reference]

Summary: Academic libraries feel the need to investigate their impact on learning, teaching and research for three reasons: 1) to demonstrate their value within the academic setting; 2) to be aware how rapid changes in innovations affect their users; and, 3) to keep abreast of the emphasis on meeting users needs. All of the conference speakers noted the difficulty of measuring impact (also noted by Oliver, 2002), and suggested measuring surrogates of impact. Several speakers also noted that negative responses to change still reflect impact. The report ended with conference attendees' conclusions of factors to consider when investigating impact, including the need to disseminate results and the difficulty of making "impact measurement an integral part of the quality improvement agenda."

Comment: Many similarities have been drawn between traditional libraries (especially at the university level) and digital libraries. While each might use different metrics and data gathering, their previous focus on measuring specific aspects (library use, effectiveness of services, patron demographics) are the same. So too, traditional and digital libraries are beginning to grapple with measuring their impact. While the primary conclusions of conference attendees were generic to all planning processes, the sub-questions, applicable to traditional university libraries, can also be considered in the context of evaluating the impact of digital libraries, especially in a university setting.

Gunn, C. (1999). "They Love It, but Do They Learn from It? Evaluating the Educational Impact of Innovations." *Higher Education Research and Development* **18**(2): p. 185-99.

Author Abstract: The SECAL (Situated Evaluation of Computer-Assisted Learning) framework offers a broadly based method for evaluating learning with technology in its many forms and implementations. Central to the framework are detailed and discipline-specific definitions of learning and corresponding descriptions of innovative study environments designed to exploit the potential of technology to support achievement of high quality learning goals. The objectives are to collect evidence of how these environments may or may not lead to effective learning and to identify what, if anything, might be done to improve the prospects. The concept of situation implies the need to evaluate contextual influences as well as how students and lecturers actually use technology. This is achieved through rich description generated from multidimensional, qualitative methods which are theoretically grounded in interpretive, critical and postmodern paradigms. The ubiquitous student evaluation of teaching systems are over-reliant on subjective data and offer little insight into pedagogical issues. The case-specific SECAL method uses objective and subjective data to assess how technology impacts on learning processes and outcomes. Broader objectives include grounded-theory development and identification of institutional influences on teaching and learning innovations. This type of evaluation is not particularly easy to conduct, but is a prerequisite to gaining academic credibility, maximizing the benefits of investment and justifying it in terms acceptable to economic-rationalist administrators. A description of the method in this article is followed by a case study illustrating its practical applications. [32 references]

Comment: This is one of the few papers included that reports a very specific instance of evaluating the impact of educational technology. Both this and Tzortzidou (2001) were considered important for their differing approaches (experimental design (Tzortzidou) vs. case study (Gunn)) to evaluating the effect of technology on

learning. Both studies advocated a long-term approach and the use of multiple metrics to evaluate various objectives.

Haertel, G. D. and B. Means (2000). Stronger Designs for Research on Educational Uses of Technology: Conclusion and Implications. Menlo Park, CA, SRI International**:** 50. Online: http://www.sri.com/policy/designkt/synthe1b.pdf

No Abstract. [27 references]
Summary: This report synthesizes key ideas presented in ten commissioned papers from research methodology experts. The papers were written to address the insufficiency of current large-scale data collections for answering questions about technology effects.. The papers make recommendations for educational technology research approaches and research funding priorities. The report authors noted that three themes appeared and reappeared in nearly all of the commissioned papers. They were:
1) T*he need for new assessment approaches to measure student learning outcomes that are not well represented on traditional standardized achievement tests.* Evaluations of technology effects suffer from the use of scores from standardized tests of content unrelated to the technology intervention and from the substitution of measures of opinion, implementation, or consumer satisfaction for measures of student learning. The authors recommend that the new learning assessments should include:
· Extended, performance tasks
· Mechanisms for students to reveal their problem-solving, to describe their rationale for proceeding through the task, and to document the steps they follow
· Opportunities to demonstrate social competencies and collaboration
· Scoring rubrics that characterize specific attributes of performance
· Scoring rubrics that can be used across tasks of varying content
· Integration with curriculum content
· Links to content and performance standards
· Content negotiated by teachers

2) *The call for careful measurement of implementation and context.* The report highlighted paper authors emphasis on the need for careful definition and measurement of the technology innovation as it is implemented.

3) *The advantages of conducting coordinated or clustered studies that share approaches, measurement instruments, and research infrastructure.* The paper authors envisioned a program of interrelated studies linked to prior research and to other studies conducted in tandem or in sequence as part of a more comprehensive research agenda. Additionally, the authors proposed the establishment of an "intermediary organizations" to provide the infrastructure to support the inter-related program of studies.

The paper authors agreed that multiple and complementary research strategies are needed to measure the implementation and impact of learning technologies. The authors suggested:
· Collection of both qualitative and quantitative data
- Assessment of a wide range of student learning, attitude, and behavioral outcome measures

· Assessment of both context and implementation, as well as the primary intervention
· Design of both small- and large-scale studies

The paper authors did not endorse a single research strategy. Instead, they recommended an eclectic approach along the lines of the following general strategies emerged:
• Multiple, Contextualized Evaluations
• Multi-Level, Longitudinal Research
• Random-Assignment Experiments
The authors then identify broad categories of research goals and circumstances with implications for the choice of methods.

The authors conclude with a summary of a proposed technology research agenda designed to address larger research questions that have not been answered by individual project-linked research or evaluation studies.
• Information System for Educational Context Measures: Examples of important contextual variables include teacher characteristics, teacher pedagogical beliefs, professional development supports, school leadership, community engagement, technology infrastructure, and the accountability system in place.
• 21st Century Skills, Indicators, and Assessments: The authors list skills sets identified by several institutions as possible to include.
• Research on Technology Use in Schools: this research program would examine the frequencies and correlates of common and emerging "naturally occurring" practices.
• Research on Teaching & Learning with Technology: this line of research would examine the student learning effects of well-defined projects or innovations involving technology.
• Research on Technology and Teacher Professional Development: this research would be conducted with the purpose of informing policy discussions around state and district accountability systems.

Comment: A nice synthesis of the thinking on current research needs and new approaches. Although it reads somewhat like a "wish-list," the recommendations are concrete enough and the context has not changed sufficiently as to render the ideas irrelevant. This is a must-read.

Kintsch, W. (1994). "Text comprehension, memory, and learning." *American Psychologist* **49**(4): 294-303.

Author Abstract: People are often able to reproduce a text quite well but are unable to use the information in the text for other purposes. Factors that help people to reproduce a text have been studied for some time. This article explores ways that enable people to learn from texts. Content overlap between a text and the reader's prior knowledge is identified as one factor, and methods are proposed to identify whether a text is suitable for readers with given background knowledge. For readers with low background knowledge, a text should be as coherent and explicit as possible to facilitate learning. However, data are presented to show that for readers with adequate background knowledge, texts with coherence gaps that stimulate constructive activities are in fact better for learning.

Summary: This article, written by the preeminent comprehension researcher in cognitive psychology, summarizes critical factors in learning from written materials as identified by comprehension research. These factors must be considered when assessing the learning impact of textual materials. Kintsch highlights the importance of distinguishing between simple memory for a text versus true understanding of a text. He discusses data from assessment techniques (free recall, problem-solving questions, card/concept sorting tasks) that he and his colleagues have used to measure text memory versus learning. Kintsch also emphasizes the importance of a student's background knowledge in determining learning performance. Students with varied levels of knowledge in a domain often learn best from different materials; the interpretation of this effect is that higher knowledge students need materials that encourage active processing whereas lower knowledge participants need more complete and coherent information.

Comment: Although this article does not explicitly address assessment of educational technology, it provides an excellent summary of what cognitive psychologists know about comprehension and highlights critical factors that must be considered in any valid assessment of learning.

Levin, J. A. and B. C. Bruce (2003). Technology as Media: A Learner Centered Perspective. What Teachers Should Know about Technology? Perspectives and Practices. Y. Zhao, Information Age Press. Online: http://faculty.ed.uiuc.edu/j-levin/levin-bruce.html

Author Abstract: The balance between learning and doing is impacted by new technologies for learning. In this paper, we explore a framework for expertise that emphasizes the power of multiple coordinated representations. We use a learner-centered taxonomy of technology uses for learning as a framework for systematically developing powerful environments for learning. This taxonomy can help us determine the most appropriate technologies for learning, given the educational goals and the available educational resources. [15 references]

Summary: The authors enumerate the costs and benefits of learning to make the point that the context in which technology is introduced will have an impact. They further note that "as the media change, then the costs of learning can change and the benefits of learning also can change." The authors then propose a "taxonomy of uses of technology for learning," based on the goals of the learner. Ultimately, the taxonomy could provide a framework to systematically think about the diversity of knowledge domains and to construct learning environments that support the development of expertise. The four elements of the taxonomy include:
- Media for inquiry
- Media for communication
- Media for construction
- Media for expression

Comment: The authors' learner-centered taxonomy of technology, initially conceived to design contexts for learning, could also provide a perspective on evaluating technology. Digital libraries incorporate several characteristics of each taxonomy element. Relating a digital library function or feature to a taxonomy element and its associated learner expertise may assist in identifying specific impacts from a broad range of possible effects.

Lin, X. (2001). "Reflective adaptation of a technology artifact: A case study of classroom change." *Cognition and instruction* **19**(4): 395-440.

Author Abstract: This case study examines the changes that took place in a 5th-grade mathematics classroom in Hong Kong when an American-made technological artifact was introduced. Using classroom observation, videotaping, and daily interviews with the teacher and students, the study documents the changes that occurred in the classroom and the factors shaping those changes. Using the artifact prompted the teacher to make changes by altering the content, the sequence of instruction, and the social arrangements for class participation. In addition, the class engaged in extensive discussions about the social implications of solving complex mathematics problems. Most of these changes, however, were made after the attempts to maintain routine instruction broke down. The study suggests that new technologies that are not compatible with the existing practice may lead to intense and sometimes profound reflections about which aspects of the individuals, their environment, and the new artifact to adapt into a new classroom structure. A significant challenge is how to support participants' role shifting and their subsequent psychological changes when new technology is introduced.

Summary: Lin argues that the following three factors must be considered when considering the interaction between educational technology, classrooms, and individual (p. 431):
1. The affordances of the artifact
2. Support and constraints offered by the local culture
3. The kinds of reflection and decisions that influence the adaptation.

Lin argues that educational technology enables, but does not dictate, changes in classroom practice and that this type of affordance provides opportunity for reflection and growth, both by teachers and by students.

Comment: This article addresses educational impact from a different perspective. Instead of focusing on individual cognition, it examines social and instructional changes (including changes in classroom practice and metacognitive reflection) that can occur when educational technology is introduced into a classroom.

Linn, M. C. and S. Hsi (2000). Computers, Teachers, Peers--Science Learning Partners. Mahwah, New Jersey, Lawrence Erlbaum Associations.

No Abstract. [203 references]
Summary: The authors drew on their 15 years of experience with the Computer as Learning Partner project to create a guide for teachers, scientists, educational researchers, technology specialists, curriculum designers and students to use to work together as partners to improve learning outcomes in science.

Section 1: Uses case studies to demonstrate the diverse responses four students have to the same science curriculum.
Section 2: Describes the development process of the Computer as Learning Partner. The authors' and developers' perspective is that computers should be viewed as learning partners rather than as tools, teachers or laboratory equipment. They present pedagogical principles around the four tenets of scaffolded knowledge integration:

- Making science accessible: connecting to what students want to know. 1) Encourage students to build on their scientific ideas as they develop pragmatic scientific principles. 2) encourage students to investigate personally relevant problems and revisit their science ideas regularly. 3) Scaffold science activities so students participate in the inquiry process.
- Making thinking visible: explaining mistakes, animating science processes and illustrating connections. 1) Model the scientific process of considering alternative explanations and diagnosing mistakes. 2) Scaffold students to explain their ideas. 3) Provide multiple, visual representations from varied media.
- Helping students learn from each other: building respectful, efficient and effective collaborations in the classroom. 1) Encourage students to listen and learn from each other. 2) Design social activities to promote productive and respectful interactions. 3) Scaffold groups to design criteria and standards. 4) Employ multiple social activity structures.
- Promoting lifelong science learning: supporting project work, reflecting on scientific ideas and revising science questions. 1) Engage students in reflecting on their own scientific ideas and on their own progress in understanding science. 2) Engage students as critics of diverse scientific information. 3) Engage students in varied, sustained science project experiences. 4) Establish a generalizable inquiry process suitable for diverse science projects.

Section 3: Examines science learning from the perspective of the four students presented in the case studies in Section 1.
Section 4: Discusses how to design effective instruction using technology as a learning partner and how classroom teachers can form partnerships to improve science education.

Comment: The authors' operationalize the idea of learning *with* technology (Reeves, 1998). Their emphasis on continuous improvement to science education, rather than looking for a quick fix or a silver bullet, is realistic and would support a long-term evaluation effort advocated by many authors in this bibliography.

Marchionini, G. (2000). "Evaluating Digital Libraries: A Longitudinal and Multifaceted View." *Library Trends* **49**(2): 304-333.

Author Abstract: The Perseus Digital Library (PDL) is one of the primary digital resources for the humanities. Under continuous development since 1987, the project has included an ongoing evaluation component that aims to understand the effects of access to digitized source materials in the humanities. A summary of the PDL genesis and current status is given and the multifaceted and longitudinal evaluation effort is described. A brief synthesis of results is provided and reflections on the evaluation along with recommendations for DL evaluation are given. [27 references]

Summary: This paper describes the evaluation efforts associated with the Perseus Digital Library (PDL), presents a synthesis of evaluation results and provides recommendations for evaluating digital libraries. The paper begins with a description of the PDL, "an evolving digital library of resources for the study of the ancient world and beyond." PDL began in 1985 with plans to digitize classic Greek texts and English translations, implemented on the Apple HyperCard platform. It has evolved to become a web-based digital library that includes thousands of text and multimedia objects, descriptive metadata and tools to search and interact with resources; it has also become a widely used resource in teaching, learning and research at all levels.

The author reports that as the project team developed the evaluation plan, they incorporated many aspects of educational evaluation but also adopted metrics and techniques from information science. Other factors that affected developing the evaluation plan were defining what the PDL was (i.e., its mission) and acknowledging the emphasis, placed by funding agencies, on the educational application of PDL materials. The team developed an open plan that provided opportunities to incorporate new information or pursue new areas of research. The evaluation plan took shape as the team crossed projects goals (access, freedom and collaboration) with objects of evaluation (learners, teachers, content and the technical system), defined a set of 94 research questions and mapped data collection techniques (observations, interviews, document analysis and learning analysis) and analysis methods onto these questions. The author describes how, as the project progressed, the focus of evaluation shifted to evaluating the impact of PDL on teaching and research, and the team made an effort to identify their users with an online questionnaire. The author also notes the types of data considered useful for the learning analysis (student reading and translating rates, assignments and syllabi characteristics) and the difficulty in obtaining some data, such as student grades. The author provides a synthesis of evaluation results and also lists pointers to results of specific yearly evaluations. The author concludes with three recommendations for conducting evaluations of digital libraries in other settings and a list of qualities that contributed to the success of PDL.
1. Evaluation efforts must explicate goals on a continuum ranging from evaluation research to product/system testing. The author describes how many formative and summative components were incorporated into the PDL evaluation plan and suggests that any evaluation plan should incorporate as many measures as feasible without relying too heavily on one specific measure.
2. Digital libraries are emergent complex systems. 2.1. Evaluation research should be longitudinal 2.2. Evaluation research should be multifaceted. The author notes that digital libraries are an "emergent phenomena" and that evaluation plans must be longitudinal, flexible and "designed to seek unexpected outcomes."
3. Integrate statistical data and narratives to assess impact as well as performance and usage. The author states, "Operational data are powerful components in a chain of inferences that address impact but the PDL evaluation illustrates the value of anecdotes and "stories" that illustrate new effects."
Digital library success is aided by: 1) clear missions; 2) strong leadership and a strong talent pool; 3) good technical vision and decisions; 4) quality content and data management; 5) giving users multiple access alternatives; and, 6) ongoing evaluation effort.

Comment: Aside from the preliminary evaluation plans described in the Borgman (2000) paper, this is the only report, with reflections, about evaluating a digital library for purposes other than technical design and development. The long-term approach advocated may not be feasible for most projects on a 2-3 year funding cycle. This argues for projects to seek diverse funding opportunities and to scale their evaluation plans to have maximum impact in the time allowed. Short funding cycles also show that there is a need for projects to be able to leverage pre-existing evaluation expertise and be able to collaborate with other projects to achieve synergistic relationships for mutual benefit.

McNabb, M., M. Hawkes, et al. (1999). *Conference Summary: Critical Issues in Evaluating the Effectiveness of Technology*. The Secretary's Conference on

Educational Technology, 1999: Evaluating the Effectiveness of Technology, Washington, D.C. Online: http://www.ed.gov/rschstat/eval/tech/techconf99/confsum.html

ERIC Abstract: The Secretary's Conference on Educational Technology: Evaluating the Effectiveness of Technology (July 12-13, 1999) noted a shift in schools' focus on technology. Where once the emphasis was on building and implementing a technology infrastructure, today it is on evaluating the effectiveness of its use in schools and classrooms. Conference participants exchanged promising evaluation strategies and techniques and considered how to respond to the many voices demanding to know technology's effects on schooling. The following seven critical issues in evaluating the effectiveness of technology in education arose as a consequence of the interaction among stakeholders: (1) The effectiveness of technology is embedded in the effectiveness of other school improvement efforts; (2) Current practices for evaluating the impact of technology in education need broadening; (3) Standardized test scores offer limited formative information with which to drive the development of a school's technology program--most schools are looking for additional means for collecting useful data for this purpose; (4) Schools must document and report their evaluation findings in ways that satisfy diverse stakeholders' need to know; (5) In order for evaluation efforts to provide stakeholders with answers to their questions about the effectiveness of technology in education, everyone must agree on a common language and standards of practice for measuring how schools achieve that end; (6) The role of teachers is crucial in evaluating the effectiveness of technology in schools, but the burden of proof is not solely theirs; and (7) Implementing an innovation in schools can result in practice running before policy. Some existing policies need to be "transformed" to match the new needs of schools using technology. [no references]

Summary: Conference participants recognized the importance context plays in evaluation, but unlike Agodini, Dynarski, et al. (2003), they thought that traditional experimental designs would not yield effective information due to the number of interacting factors. The emphasis of other papers listed here has been to disseminate results of a particular study; the papers have been less concerned with using evaluation results to affect policy or administrative decisions. The summary indicates that participants were acutely aware of stakeholders involved in evaluation, and discussion around many of the critical issues was driven by disseminating and acting on the results. Participants did not place an emphasis on informing a R&D cycle for improving technology. Rather, they expected to use the finished technology product and to have evaluation inform and improve their use and their students' learning.

Comment: The abstract reflects the conference summary. For further reading, there are 13 commissioned white papers available on the conference website (http://www.ed.gov/Technology/TechConf/1999/). Similar to the *2000 Secretary's Conference on Educational Technology* (Reed, 2000), the critical issues identified represent the thoughts and experiences of "front-line" teachers.

Murphy, R. F., W. R. Penuel, et al. (2001). E-DESK: A Review of Recent Evidence On the Effectiveness of Discrete Educational Software. Menlo Park, Ca., SRI International**:** 119 pages. Online: http://www.ctl.sri.com/publications/downloads/Task3_FinalReport3.pdf

No Abstract. [106 references]
Summary: This report synthesizes results from studies conducted between 1993 and 2000. It takes a systematically selective approach by including only research that "meets strong standards of methodological quality and relevance to the question of discrete educational software's [DES] impact on reading and mathematics achievement." Types of DES have traditionally included integrated learning systems, computer-based instruction and computer-assisted instruction along with newer software programs available on CD-ROM or the Internet. The authors note earlier syntheses produced conflicting results about the positive association between software use and student achievement. Their research shows that "the current research base provides little guidance on program effectiveness for policy-makers and practitioners because of the scarcity of rigorous studies of program impacts." The authors formulated three findings from their research:
1) The poor overall quality of the current state of effectiveness research restricts researchers' ability to learn from the experiences of others and limits the development of a knowledge base that will help inform the work of decision-makers, practitioners, and designers.
2) A positive association exists between DES use and achievement in math and reading; however, the ability to generalize from these results is limited.
3) The failure of studies to report basic information on effect size and implementation is a major barrier to developing new knowledge in the field.
To improve the quality of evaluation research, the authors recommend that:
1) Researchers in the field establish clear guidelines for evaluation designs that set out minimum requirements for scientifically defensible research using the health care field as an exemplar.
2) Peer-reviewed journals place stricter reporting standards on accepted manuscripts.
3) Government entities adopt stricter reporting standards for what deliverables should include.
4) More emphasis be given within quasi-experimental evaluation designs to understanding *why* programs are effective and *for whom do they work best*.
5) Data on program outcomes and on the nature of the treatment be collected within the same studies.
6) Researchers in the field consider adopting the emerging standards in the social sciences for reporting data in ways that allow for comparison of research findings across studies.

Comment: This report is useful because it provides clear ideas for how future evaluation research should be structured in order to fill gaps in the existing knowledge base. Unfortunately, studies examined for this bibliography seem not to have incorporated the authors' suggestions.

Oliver, M. and J. Harvey (2002). "What Does 'Impact' Mean in the Evaluation of Learning Technology?" *Educational Technology & Society* **5**(3): 18-26. Online: http://ifets.ieee.org/periodical/vol_3_2002/v_3_2002.html

Author Abstract: Whilst many projects in Higher Education are expected to demonstrate their impact, quite what this requirement means is often left unspecified. This paper draws on the experiences of the EFFECTS project in an attempt to illuminate this issue. The EFFECTS evaluation framework is used to structure this discussion, which explores the complexities associated with identifying

impact in terms of student learning, changes in practice for academics, changes within an organisation and national development. Common themes arising from these areas are then identified. Importantly, while practical issues are considered, the purpose of this discussion is not to 'solve' this complexity – instead, its purpose is to recognise it, and to consider the implications for evaluators of working in such settings. [25 references]

Summary: This paper analyzes the concept of impact as it applies in the context of evaluating a learning technology project, specifically, the EFFECTS project. The authors describe levels of impact:
1) Impact on students: The authors note that most evaluations attempt to measure learning or behavioral changes and contend that "it is impossible to measure the impact of initiatives on learning without taking a theoretical stance on what learning is." Some approaches to measuring learning are through assessment "as a proxy for learning," process and direct observation. Demonstrating a change in skills is as difficult as measuring learning. However, changes in student perceptions and attitudes, while relatively easy to investigate, "tend to be viewed as a weak form of impact."
2) Impact on academics: Many of the issues that surround measuring the impact of technology on students also apply to academics. The most common areas investigated include skills and attitudes.
3) Institutional impact: The authors report the issues associated with implementing the EFFECTS project in institutions. Namely, through interviews, the evaluators gathered data on areas staff perceived as having been affected; they did not gauge the extent of impact, nor did they question the credibility of the interviewees.
4) National impact: Again, the authors draw on the issues related to measuring the national impact of EFFECTS. In this case, awareness was an indicator of impact, though they noted the pitfall of "publications being taken as a proxy for awareness." Contact lists were also used to note awareness, though there were drawbacks to this measure as well.
The authors conclude that an "investigation cannot be perceived as a value-free enquiry into truth." Few existing measures adequately and validly measure impact without being influenced by, among other things, the agenda of the evaluator, or funding agency, and the need to present a positive image by those being studied.

Comment: The paper does seem quite cynical (authors' term) in concluding that so much of the previous work done to measure "impact" doesn't really provide an accurate measure. The paper does a good job of identifying the complexities associated with identifying impact at various levels. A must read.

Penuel, B. W. and B. Means (in press). "Implementation variation and fidelity in an inquiry science program: An analysis of GLOBE data reporting patterns." *Journal of Research in Science Teaching*: 58 pages. Online: http://ctl.sri.com/publications/displayPublication.jsp?ID=129

Author Abstract: This paper examines variation in patterns in the enactment of a large-scale K-12 science inquiry program. Student data reports in the Global Learning and Observations to Benefit the Environment (GLOBE) program provide a useful measure of implementation, since key design elements in the program are student collection and reporting of local environmental data. We examined associations between teachers' responses to survey items to patterns in GLOBE data

reporting to develop hypotheses about important contextual factors that are related to program implementation. Implications for the study of science inquiry programs are discussed. [48 references]

Summary: GLOBE is an international environmental science and science education program focused on improving student understanding of science by involving them in the collection of data for real scientific investigations. The purpose of this evaluation research was to "analyze how well the assumptions underpinning the GLOBE design stood up to data on GLOBE program implementation." Before describing the methodology and results of the research, the authors identify previous research and ideas that informed their work. Their review indicated that researchers study the variations in the implementation of educational innovations to "understand the limits of a program's applicability or flexibility and in explaining within-innovation variations in effectiveness (Lipsey & Cordray, 2000) and to identify competing hypotheses for observed impacts (Schiller, 2001) and possible flaws in the assumptions that underlie the program design (see Goodson, Layzer, St. Pierre, Bernstein, & Lopez, 2000)." They noted that researchers measure implementation fidelity (the extent to which teachers enact innovations in ways that either follow designers' intentions or that replicate practices developed elsewhere (Loucks, 1983)), though other researchers note that this misses the role teachers play in adapting innovations to a certain context. They observed that "[f]ew studies to date, however, provide a context for studying innovations where both variation [in implementation] and fidelity are important dimensions of implementation to designers." As the report describes, the researchers are able to measure the GLOBE program for variation and fidelity in implementation by reviewing information from a data archive and teacher reports.

Comment: The structure of the study, and its results, are applicable primarily to similarly structured science inquiry programs. More relevant to evaluating the educational impact of digital libraries is the authors' debate of the merits of evaluating "implementation fidelity" or "adaptation in implementation." This will one factor to consider when designing a research study.

Quellmalz, E. and A. M. Haydel (2003). *Using cognitive analysis to study the validities of science inquiry assessments*. American Educational Research Association Annual Meeting, New Orleans, LA.

No abstract. [27 references]
Summary: This paper describes a study to measure the cognitive validity of tests that accurately represent students' science learning. Its purpose was to compare the cognitive components of science performance elicited by the different formats and to document a cognitive analysis methodology. The authors note that, "Cognitive validity has been identified... as a form of construct validity that provides evidence of the declarative knowledge and procedural knowledge elicited by assessments designed to test concepts and skills. Evidence for cognitive validity is based on the correspondence of what tests intend to measure, what cognitive operations are actually elicited by the tests, and the scores students receive for their performance" (p. 1). The study addressed issues in identifying the types of item content and format most likely to yield evidence of students' inquiry strategies, development of the protocol procedures that would elicit students' reasoning through investigation strategies, types of follow-up probes that would elicit further information about

inquiry and test perceptions, and the coding schemes that would yield useful information about students' inquiry (p. 3). The authors concluded that: 1) Multiple-choice items on the reference exams were less likely to tap inquiry or complex conceptual understanding; 2) Constructed-response items seemed to be tapping a fuller range of inquiry and conceptual understanding, though they seemed to provide little insightful information when items were non-inquiry; and, 3) Performance assessments seemed suited to tap the full range of inquiry and conceptual understanding, depending on the design of the items within the performance assessments. Ultimately, the authors conclude that in order to determine the level of students' knowledge of science content and inquiry, multiple item formats are needed (p. 7).

Comment: Although this study does not analyze assessments that measure the impact of technology on science learning, the methods used could provide a model for assessing the impact of technology. The study demonstrates what other authors in this bibliography have advocated: the use of multiple measures to fully assess student knowledge and inquiry in science learning.

Reed, D. (2000). *Conference Summary: How should the value of technology to learning be measured?* The Secretary's Conference on Educational Technology, 2000: Measuring Impacts and Shaping the Future, Alexandria, VA. Online: http://www.ed.gov/rschstat/eval/tech/techconf00/report.html

No Abstract. [no references]
Summary: This report summarizes key questions, topics and conclusions from the second of two conferences, the first of which was held in 1999. To guide discussion, organizers posed seven questions and commissioned x [get number] whitepapers which participants deliberated in breakout sessions. Participants were a mix of k-12 teachers, educational researchers and industry representatives. The report reflects the conference structure by listing the seven questions and then providing supporting quotes and recommendations from whitepapers and breakout discussions. In the course of research, there were no references to further workshops after this.
Question 1: What constitutes the effective use of technology in learning? What value does technology bring to learning? Participants noted that the context in which technology is used is key to its effectiveness. However, teachers described how technology had not been effectively integrated into the curriculum resulting in their students viewing inquiry-based learning [facilitated by technology] as fun, but not relevant, since student grades reflected standardized tests based on textbooks.
Question #2: Will we recognize effective uses of technology when we see them? Participants responded with three observations: a) Indicators of effective uses have yet to be determined as teachers feel pulled to teach skills necessary to succeed in a digital world while being pushed to also prepare students for standardized tests. b) Schools are struggling to find common ground between the traditional assessment tools and new, almost non-existent, tools that assess skills and knowledge learned with technology. c) Schools need to become high performance, high technology systems where "system" applies as much to "learning cultures that are open to innovation…" as to networks.
Question #3: What uses of learning technology does the public value? Participants recognized that broader social and work-related requirements often dictate the value of skills or knowledge, which is then incorporated into education standards. They

acknowledged that teachers should be more active in educating the public about what should be taught in schools.
Question #4: What conditions must be in place in schools to ensure effective technology use? Participants developed a laundry list of conditions. Highlights included teachers being proficient with technology tools in order to use them effectively, teachers having the capacity to test higher-order thinking skills and assessments reflecting new technology standards.
Question #5: How can we successfully gauge and report progress with technology at the educator proficiency and system-capacity levels, as well as at the student performance level? Participants expressed a need for technology-based evaluation tools, and one suggested setting specific standards, publicizing progress better and making data-driven decisions.
Question #6: What is the policy roadmap that would build the capacity of communities and schools to move toward more effective uses of technology in schools? A specific roadmap was not developed during the conference, but teachers acknowledged the need to develop relationships with policymakers while cautioning that the achievements policymakers would like to see do not happen quickly.
Question #7: What can we learn from business and industry? Many participants related positive experiences with school-business partnerships. They noted that the partnerships resulted in better access to technology and another perspective about the end result of teaching and learning with technology.

Comment: While this summary is pitched at a high level, the perspective of teachers on the "front line" is important to remember when developing an evaluation strategy. The context (in this case, k-12 classrooms with aging technology lead by teachers pressured to teach to standardized tests) in which technology, including digital libraries, is used will, in part, determine its impact.

Reeves, T. (1998). The Impact of Media and Technology in Schools. A Research Report prepared for The Bertelsmann Foundation. Athens, Ga, The University of Georgia. Online: http://www.athensacademy.org/instruct/media_tech/reeves0.html

No Abstract. [146 references]

Summary: This report summarizes the evidence for the effectiveness and impact of media and technology in K-12 schools around the world. Reeves defines media (the symbol systems that teachers and students use to represent knowledge) and technology (the tools that allow teachers and students to share their knowledge representations with others). He distinguishes between learning "from" and "with" media and technology. The learning "from" approach entails media and technology functioning as a tutor to communicate content to learners with the assumption they will learn from the communication. Primary technologies used in this approach are the television and computer. The value of learning "from" media and technology lies in their capacity to motivate students, increase equity of access and reduce the time needed to accomplish a give set of objectives. In the learning "with" approach media and technology cognitive tools, such as databases, online collaborative software or multimedia construction software, are given directly to learners to facilitate critical thinking and higher order learning. Value derived from learning "with" cognitive tools occurs when the tools are applied in a constructivist learning environment and when the tasks requiring cognitive tools are situated in realistic contexts with personally meaningful results. Reeves also noted that cognitive tools have two kinds of

cognitive effects: "those which are *with* the technology in terms of intellectual partnerships and those that are *of* the technology in terms of the cognitive residue that remains after the tools are used." There are four conclusions to Reeves' research:
1. Media and technology are effective in schools to learn from and with.
2. Longitudinal studies show that pedagogical innovations and positive learning results do eventually emerge from the infusion of media and technology into schools, but the process takes longer than most people imagine.
3. Pedagogy that is most influential on learning, not media or technology. Large investments in time and support for teachers are critical if the adoption of constructivist pedagogies accompany the infusion of media and technology.
4. There is a need for long-term, intensive research focused on the mission of improving teaching and learning through media and technology with the purpose of improving, not proving. Due to the complex nature of human learning, there may be no generalizable best approach to using media and technology in schools.

Comment: Reeves' report nicely summarizes ideas expressed by other authors in this bibliography, and while digital libraries were not specifically mentioned, his conclusions are just as applicable to that technology.

Renninger, K. A., L. Farra, et al. (2000). *The Impact of The Math Forum's Problem(s) of the Week on Students' Mathematical Thinking*. International Conference of the Learning Sciences 2000, The Regents of the University of Michigan.

No Abstract. [2 references]
Summary: This study reports the effort to evaluate the impact of The Math Forum's Problem(s) of the Week (PoW) on students' mathematical thinking. Skills associated with the mathematical thinking include the use of strategies and reflection as part of the process of problem-solving. The authors describe the structure and purpose of the PoWs and then relate the methodology. The authors drew a random sample of PoWs from the MathForum archives, created a sample of 40 students and a control group and reviewed problems over a 10-month time span, identifying a Time 1 and a Time 2 for comparison. They rated the PoWs on problem difficulty, mathematical thinking and on the narrative descriptions of students' mathematical thinking. The authors found that student levels of connectedness, strategy use, and autonomy increased from Time 1 to Time 2. The authors also examined change, "where the independent variables were student preparation and identification as a person who revised and resubmitted (or did not)." The authors conclude that 1) mathematical thinking of students in elementary, middle school, and high school is enhanced through work with the PoWs; 2) students who are weak with respect to their ability to connect to and use strategies to solve problems, and/or need structured facilitation (lack autonomy) at the outset of their work with the PoW, make even more dramatic gains than do their counterparts who have stronger preparation; and, 3) based on narrative data, students are most likely to use feedback that specifically indicates what they need to do when they revise and resubmit their work.

Comment: This study provides a model for a small-scale evaluation of the effectiveness of a particular intervention which is facilitated by technology.

Rogers, J. D. and B. Bozeman (2001). ""Knowledge Value Alliances": An Alternative to the R&D Project Focus in Evaluation." *Science, Technology and Human Values* **26**(1): 42 pages. Online: http://www.rvm.gatech.edu/paperfiles/00-01/juan-kva.doc

Author Abstract: The question of what the relevant entities or units of analysis for studying the dynamics of R&D are is central not only for adequate characterizations of the system of scientific and technological knowledge production but also for determining the correct focus for evaluation of R&D activities. Typically, R&D performance evaluations have focused not only on the wrong thing, but have looked in the wrong place. Most evaluations have been project or program based. Often this focus is misleading.

In this paper, we present a "knowledge value" framework as an alternative focus for understanding and evaluating scientific and technical work. This framework consists of two core concepts: the Knowledge Value Collective (KVC) and the Knowledge Value Alliance (KVA). Based on the analysis of 28 case studies of research activities we present a typology of KVAs and conclude that they are a better object of evaluation than discipline based projects. [36 references]

Summary: Comprehensive abstract; no summary needed.

Comment: This paper presents an alternate framework for performing evaluations with a shift away from the project-focused evaluations that are typical of digital library development efforts.

Saracevic, T. (2000). "Digital library evaluation: Toward an evolution of concepts." *Library Trends* **49**(2): 350-369. Online: http://scils.rutgers.edu/~tefko/LibraryTrends2000.pdf

Author Abstract: While there were many efforts in the research and practices of digital libraries, evaluation was not a conspicuous activity. It is well recognized that digital library evaluation is a complex and difficult undertaking. Challenges facing digital library evaluation are enumerated. A conceptual framework for evaluation is suggested. A review of evaluation efforts in research and practice concentrates on derivation of criteria used in evaluation. Essential requirements for evaluation are stated. Discussed are constructs, context, and criteria of digital libraries: What should we evaluate? For what purpose do we evaluate? Who should evaluate? At what level do we evaluate! Upon what criteria do we evaluate? In addition, included are suggestions for adaptation of criteria from related activities. The article is considered as a part of the evolution of concepts for digital library evaluation. [27 references]

Summary: The author notes there are two ends of the digital library development continuum, research and practice. Both face various barriers to evaluation due, in part, to the culture from which they have evolved and their ultimate use. He reviews results of the research community's efforts by summarizing several DLI-1 projects and their evaluation efforts (see Borgman, 2000). His main point from this review is: "User studies, while useful for understanding how people use systems, by themselves are not evaluation even thought they may have evaluative implications and they provide important criteria that can be used in evaluation." He then reviews results from the practice community's efforts, including the Perseus Digital Library

(Marchionini, 2000) and notes that evaluation efforts, driven by the need to develop operational digital libraries, included criteria such as efficiency and economic impacts. The author also notes traditional libraries' history of and contributions to evaluation.
The author takes a systems (as opposed to an ethnographic, sociological, economic) approach in deciding why to do an evaluation. He views the digital library as a system and reviews several definitions of a digital library in order to derive a list of elements that could be evaluated. He notes those elements occur in contexts which affect their performance and divides the contexts into user centered (social level, institutional, individual, interface) and system centered (engineering, processing, content). He then lists criteria from other fields of practice (traditional libraries, information retrieval, HCI) that could influence digital library evaluation.

Comment: Article presents one approach to evaluating digital libraries. It's interesting to see the logical progression of Saracevic's thoughts as he follows a systems-approach to evaluation to its conclusion. His definitions of digital libraries are worth reviewing, as are the elements for evaluation he identified. Also good to keep in mind is his emphasis that using one metric does not substitute for a complete evaluation. However, I think his strictly systems approach could be limiting in its emphasis on evaluating the more technical aspects of a digital library. Marchionini's (2000) paper operationalizes some of the processes outlined here in a working digital library environment. Also, compare this process with that outlined in Agodini (2003) for an educational approach that incorporates evaluating technology in general with the goal of measuring the effect of technology on learning, not just measuring user behavior through direct interactions with technology or digital libraries.

Sivin-Kachala, J. and E. Bialo (2000). 2000 research report on the effectiveness of technology in schools (7th ed.). Washington D.C., Software and Information Industry Association**:** 136 pages. Online: www.siia.net

No Abstract. [327 references]
Summary: This report summarizes research from 311 studies (out of 3,500 reviewed) on educational technology used in k-12 and higher education from the late 1980's through 2000. The primary audiences for this report are software developers and publishers and educators who will purchase and incorporate software into their curriculum; however, the report is also a useful research tool in that its conclusions focus on operational aspects of successful design and implementation of educational software. The report does not include studies of digital libraries or their effectiveness, and only includes one study about evaluating the online catalog interface in a traditional library (Eliasen, McKinstry, Fraser and Babbitt (1997)). The report is divided into three sections:
- Effects of technology on student achievement
- Effects of technology on student self-concept and attitude about learning
- Effects of technology on interactions involving educators and students in the learning environment.

Comment: This report complements Murphy, Penuel, et al. (2002) and Reeves (1998) as a comprehensive survey of education technology.

Taylor, J. (in press). *A Task-centred Approach to Evaluating a Mobile Learning Environment for Pedagogical Soundness*. MLEARN2003, London, UK. Online: http://kn.open.ac.uk/public/document.cfm?documentid=3102

Author Abstract: The focus of this paper is how to evaluate the pedagogical soundness of a mobile learning environment in which many users (both teachers and learners) may not have previously encountered mobile technology, so may be uncertain as to how best to deploy it to achieve their goals. Drawing on concepts from Activity Theory and the socio-cognitive engineering method described by Sharples, (2000) an approach is described which enables an enriched view of users' current and future activities, which in turn will allow us to understand the range of actions and opportunities for mobile learners, and seek ways of extending this range to support what learners want to do - even if they themselves do not yet know what that is. [7 references]

Summary: This paper describes the approach used to evaluate the European-led research and development project MOBIlearn. The author notes that there are "no existing comprehensive frameworks for broader formative evaluation in the mobile environment, largely due to its novelty." Since this project has just begun, the author describes how new ways of engaging in learning are being introduced at the same time as new ways of evaluating technical and pedagogical effectiveness. The researchers will use a two-step socio-cognitive engineering method of system design to analyze users' activities and then design the new technology to integrate with the users' current environment. This will allow evaluators to analyze learners in their context and understand the nature of their learning tasks.

Comment: Since the MOBIlearn project is relatively new, there are no results to report from the evaluation. However, this project serves as a model, similar to other digital library projects (Borgman, C., A. Gilliland-Swetland, et al. 2000; Marchionini, 2000), for incorporating evaluation into every phase of project design and development from the very beginning.

Tzortzidou, S. and G. Hassapis (2001). "Assessment of the Reading Skill Improvement in the Computer-Assisted Teaching of a Foreign Language." *Education and Information Technologies* **6**(3): 177-191.

Author Abstract: This article reports the findings of a case study on the influence that the computer-assisted teaching of a foreign language has on the reading skill of elementary school pupils. In fact, it has been investigated the degree to which abilities inferred from the reading process are improved by delivering a typical introductory course of the French language to a Greek speaking target group, based on the use of the generic functions that most of the computer-assisted instructional tools provide. The study was focused on the abilities of perception, information retrieval and concentration and was conducted over a period of three years. It was based on the design and execution of pupil assessment tests on a control and an experimental group per year. At each year, the control group was taught the subject in the traditional textbook-based approach and the experimental group by a computer-assisted approach. The presentation of the textbook teaching material was enhanced with features that only the use of the generic functions of a computer-assisted teaching could offer. An extensive statistical evaluation of the results has

disclosed that the use of the computer had definitely a positive influence on the abilities of perception and information retrieval. [46 references]

Comment: This paper provides a specific example of assessing various student abilities affected by technology. It took place in a local k-12 environment using an experimental design and null hypothesis statistical methods to evaluate results. The drill-and-repeat nature of language learning is conducive to technology intervention, and the abilities under study, perception, information retrieval and concentration, are important in any discipline. In fact, they could be considered some of the abilities necessary to undertake tasks requiring higher order thinking skills, which other evaluations of the impact of technology are trying to measure (Baylor 2002). However, the authors do not interpret their findings to apply to a broader context. A key element of the experiment design, also noted by Marchionini (2002) was that longitudinal studies be used to collect data over a long period of time.

Wiley, J. and J. F. Voss (1999). "Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text." *Journal of Educational Psychology* **91**(2): 1-11.

Author Abstract: In 2 experiments, understanding of historical subject matter was enhanced when students acted as historians and constructed their own models of an historical event. Providing students with information in a web site with multiple sources instead of a textbook chapter, and instructing them to write arguments instead of narratives, summaries, or explanations, produced the most integrated and causal essays with the most transformation from the original sources. Better performance on inference and analogy tasks provided converging evidence that students who wrote arguments from the web sources gained a better understanding than other students. A second experiment replicated the advantage of argument writing even when information was presented as an argument.

Summary: Wiley and Voss' research used an impoverished web environment (for example, there were no direct links between sources) to test student use of multiple documents during learning versus student learning from a textbook. All materials reflected consistent viewpoints, rather than representing different sides of a controversial issue.
Wiley and Voss instructed students to produce one type of written assignment: a summary, a narrative, an explanation, or an argument. The researchers analyzed written products for the proportion of borrowed sentences, transformed sentences, and for the number and type of connections used (e.g., causal connections). Characteristics of the explanation assignments were quite similar to those of the argument assignments, but a statistical analysis of the quality of explanations did not reach significance over the quality of narratives or summaries, whereas the analysis of argument quality did. (The advantage of argument format in these experiments may be interpreted as suggesting that more valuable assignments reflect authentic activities in a domain.) Wiley and Voss also tested several learning measures:
1. Sentence verification (measures memory for the materials).
2. Inference verification (measures learning from the materials).
3. Principle Identification Task (measures learning from the materials).

Students who read from a textbook were found to have better text memory than students using the web sources, but demonstrated worse learning performance than students who used web sources in addition to writing arguments about the materials. Wiley and Voss reach two main conclusions:

1. Learning situations in which multiple documents are used facilitate cognitive processes involved in developing true understanding of texts, rather than just memorization of content.
2. Construction of reasoned arguments seems to promote integration of learning materials and leads to better understanding of concepts and underlying principles.

Wiley and Voss suggest that their findings indicate that web-learning situations offer a rich research area but that specific assignments (such as the requirement to write an argument) may be needed to support student learning. They also argue that future research should address how students use a multi-source learning environment to better understand its learning benefits. Finally, they point out that this research further highlights the difference between memory for information and deeper learning and argue that researchers must be careful to measure student performance with appropriate measures.

Comment: Although this research doesn't use a scientific domain, it does suggest that the multiple sources available in digital libraries could offer important learning benefits. Further, it provides some ideas about how to measure memory for and learning from materials.

Acknowledgements

The authors wish to thank the review committee, who volunteered their time to provide comments and suggest resources during the development of this document: Howard Burrows, Lillian Cassel, Jim Dietz, Jim Dorward, Ellen Iverson, Casey Jones, Mary Marlino, Flora McMartin and Tammy Sumner.

Methodology

This annotated bibliography was developed and compiled by Sarah Giersch (providing research from the digital library, library science, education and education policy fields) with contributions from Tom Reeves (providing research from the education field) and Kirsten Butcher (providing research from the cognitive psychology field).

To produce the final document, Giersch developed a citation list, with contributions from the other authors. This list was submitted to the review committee, who provided suggestions for further research. After three iterations of the citation list, Giersch wrote the annotations and analysis. The committee commented on a final draft, and Giersch incorporated any final suggestions before submitting the finished bibliography.

Giersch's research was supported by resources available at the University of North Carolina at Chapel Hill (UNC), including print and electronic databases, indexes and journals of peer-reviewed material. The preliminary citation list was compiled using a systematic, but iterative, approach to examining UNC's resources. This involved identifying potentially valuable databases and peer-reviewed journals and assigning a priority to them. Valuations were based on the source's name and description and previous use (if any). Only high priority resources were searched for this deliverable. See Appendix A for a table of resources and search terms.

However, serendipity, "known–person or -resource" searching, using the search-engine *Google* and refining search terms, especially after comments from the review committee, often proved just as useful at uncovering relevant resources as did systematic search strategies. As a result, this deliverable also contains citations gleaned from article or report "References" sections, citation and Web searches and reviewing potential articles in context of the whole journal.

As noted in the introduction, this annotated bibliography was commissioned to support the NSDL Evaluation Workshop. Resources were selected based on the ideas and perspectives the authors, and review committee, deemed necessary for attendees to be aware of when participating in the workshop. Therefore, resources included in the final document are not necessarily representative of a particular discipline. Given the specific purpose of the workshop, though, the resources often represent the *only* current thinking on the topic in each field.

Appendix A: Databases, Journal & Search Terms

The following tables list the databases and journals systematically searched during this literature review and the search terms used separately or combined in the search process.

Key

- Priority: 1=high, search first; 2=mid, search second; 3=low, search if time permits
- Type: DB=Database; JOU=Journal
- Subject: CompSci=Computer Science; Edu=Education; ILS=Library Science; Psy=Psychology/Sociology

Priority	Type	Subject	Title
1	DB	CompSci	ACM Digital Library Via ACM - Contains Full Text
1	DB	CompSci	Books 24x7 Access to this resource has been cancelled, effective 2003-05-30 - Contains Full Text
1	JOU	CompSci	ACM Publications
1	JOU	CompSci	Annual Conference on Innovations and Technology in Computer Science Education. Proceedings (ITiCSE) Holdings: 1996 - 2002
1	JOU	CompSci	Communications of the ACM
1	JOU	CompSci	Computers & Education Holdings: 5 years ago -
1	DB	Edu	EdResearch Online - Contains Full Text
1	DB	Edu	Education Index Via OCLC FirstSearch
1	DB	Edu	ERIC Via EBSCOHost
1	DB	Edu	ESubscribe - Contains Full Text
1	DB	Edu	PsycInfo Via Silverplatter
1	DB	Edu	Social Work Abstracts Via Silverplatter
1	DB	Edu	Sociological Abstracts
1	JOU	Edu	British Journal of Educational Technology Holdings: 1997 - **see all of 34 (2) Mar03 and 34(3) June03 (on web)**
1	JOU	Edu	Computers & Education Holdings: 5 years ago -
1	JOU	Edu	Education and Information Technologies Holdings: 1997 -
1	JOU	Edu	Educause Review Holdings: Vol. 33 (Jul./Aug. 1998)-
1	JOU	Edu	Information Technology in Childhood Education Annual (ITCE)
1	JOU	Edu	Instructional Science Holdings: 1997 -
1	JOU	Edu	International Journal of Qualitative Studies in Education Holdings: Vol. 10 (1997)-
1	JOU	Edu	Internet and Higher Education Holdings: Vol. 1 (1998)-
1	JOU	Edu	Journal of Asynchronous Learning Networks Holdings: Volume 1, Issue 1-
1	JOU	Edu	Journal of Computer Assisted Learning Holdings: 3/1/1999 -
1	JOU	Edu	Journal of Education for Teaching: International Research and Pedagogy Holdings: Vol. 21 (1995)-
1	JOU	Edu	Journal of Higher Education Holdings: Vol. 1, no. 1 (Jan. 1930)-
1	JOU	Edu	Journal of Science Education and Technology Holdings: 1997 -
1	JOU	Edu	Journal of Science Teacher Education Holdings: 1997 -
1	JOU	Edu	Learning and Instruction Holdings: 5 years ago -
1	JOU	Edu	Learning Environments Research Holdings: 1998 -
1	JOU	Edu	Meridian: A Middle School Computer Technologies Journal Holdings: Vol. 1, issue 1 (Jan. 1998)-
1	JOU	Edu	Pedagogy Holdings: 2001 -
1	JOU	Edu	Research in Higher Education Holdings: 1998 -
1	JOU	Edu	Studies in Educational Evaluation Holdings: 5 years ago -
1	JOU	Edu	Teaching and Teacher Education Holdings: 5 years ago -
1	JOU	Edu	Technology Source Holdings: Mar. 1997
1	JOU	Edu	The Journal of General Education Holdings: 1999 -
1	JOU	Edu	Journal of Technology Education
1	DB	ILS	ARL Directory of Scholarly Electronic Journals and Academic Discussion Lists - Contains Full Text
1	DB	ILS	Library and Information Science Abstracts (LISA)
1	DB	ILS	Library Literature & Information Science Via Wilson Web
1	JOU	ILS	Behaviour and Information Technology Holdings: Vol. 15 (1996)-

1	JOU	ILS	Journal of the American Society for Information Science Holdings: Vol. 37-Vol. 51 (1986-2000)
1	JOU	ILS	Journal of the American Society for Information Science and Technology Holdings: 2000 -
1	JOU	ILS	Library & Information Science Research Holdings: 1999 -
1	JOU	ILS	Library Hi Tech Holdings: 1997 -
1	JOU	ILS	Portal: Libraries and the Academy Holdings: 2001 -
1	DB	Psy	CogPrints Electronic Archive - Contains Full Text
1	JOU	Psy	Cognition Technology & Work Holdings: 1999 -
2	DB	Edu	Bibliography of Education Theses in Australia
2	DB	Edu	Science's Next Wave - Contains Full Text
2	JOU	Edu	Australian Journal of Educational Technology Holdings: Vol. 1, no. 1 (winter 1985)-
2	JOU	Edu	British Journal of Educational Studies Holdings: 1997 -
2	JOU	Edu	Educational Psychologist Holdings: 2000 -
2	JOU	Edu	Educational Psychology Review Holdings: 1998 -
2	JOU	Edu	International Journal of Educational Research Holdings: 5 years ago -
2	JOU	Edu	Journal of Educational Change Holdings: 2000 -
2	JOU	Edu	Journal of Educational Sociology Holdings: Vol. 1, no. 1 (Sept. 1927)-
2	JOU	Edu	Journal of Philosophy of Education Holdings: 1997 -
2	JOU	Edu	Journal of Research in Science Teaching Holdings: 1996 -
2	JOU	Edu	Science & Education Holdings: 1997 -
2	JOU	Edu	Science Education Holdings: 1996 -
2	JOU	Edu	Social Psychology of Education Holdings: Vol. 2 (1997)-
2	JOU	ILS	First Monday Holdings: Vol. 1, issue 1 (May 1996)-
2	JOU	ILS	World Wide Web Holdings: Vol. 1 (1998)-
2	JOU	Psy	Applied Cognitive Psychology Holdings: 1996 -
3	DB	CompSci	ScienceDirect Via Elsevier - Contains Full Text
3	DB	CompSci	Today's Science on File - Contains Full Text
3	DB	Edu	Gateway to Educational Materials - Contains Full Text
3	DB	Edu	PolicyFile (Public Policy Research and Analysis) - Contains Full Text
3	JOU	Edu	Brookings Papers on Education Policy Holdings: 2000 -
3	JOU	Edu	Cognitive Development Holdings: 1/1/1999 - Note(s): Full text before 1995 is unavailable.
3	JOU	Edu	Curriculum Inquiry Holdings: 1997 -
3	JOU	Edu	Early Childhood Research Quarterly Holdings: 1999 -
3	JOU	Edu	Economics of Education Review Holdings: 5 years ago - Note(s): Full text before 1995 is unavailable
3	JOU	Edu	European Journal of Education Holdings: 2000 -
3	JOU	Edu	Higher Education Holdings: 1997 -
3	JOU	Edu	Higher Education in Europe Holdings: Vol. 25 (2000)-
3	JOU	Edu	Higher Education Quarterly Holdings: 1997 -
3	JOU	Edu	International Journal of Computers for Mathematical Learning Holdings: Vol. 2, no. 1 (1997)-
3	JOU	Edu	International Journal of Educational Development Holdings: 5 years ago -
3	JOU	Edu	International Journal of Science Education Holdings: 1999 -
3	JOU	Edu	Journal of Education Policy Holdings: 1999 -
3	DB	Psy	Health and Psychosocial Instruments (HAPI) Via UNCLE

Search Terms to Use
education(al) impact
learning impact
digital library
education(al) technology
learning technology
"educational impact" and "digital library"
"educational impact" and "digital libraries"
"educational impact" and "learning technology"
"educational impact" and "learning technologies"
"learning impact" and "digital library"
"learning impact" and "digital libraries"
"learning impact" and "learning technology"
"learning impact" and "learning technologies"
technology impact research vs. educational impact research - Allison Druin
Google - effect "education technology"
Google - impact "education technology"

*Google - impact "learning technology"
*Google - effect "learning technology"
Google impact "digital library" learning
Google effect "digital library" learning
ERIC - outcomes assessment
ERIC - Technology uses in education
ERIC - Technology Integration
ERIC - query : (((learn ADJ technolog*) AND (evaluat*))) AND (1998< Publication_Date <2004)
ERIC - computer assisted instruction
ERIC - educational quality
ERIC - ((evaluat NEAR/3 impact))