Research questions
The classroom management systems developed by Brown and Campione and Scardamalia and Bereiter raise a number of interesting questions, especially in considering extensions of these approaches to other populations of learners in professional or other advanced training settings.

(1) With respect to the role of the teacher: what is the nature of teacher support? how much expert intervention/guidance is needed? and how does this change depending on the composition of the group? In a classroom setting, especially where students are in the initial stages of acquiring new domain knowledge, unguided exploration has not been shown to work well (cf. Helmke & Weinert, 1995; Weinert, 1995).

Students in CSILE largely set their own curriculum of inquiry about "interesting facts" and assume the role of providing feedback to each other via commentary on notes, while the teacher models expert learning strategies. Classroom management techniques structure activities, such that different groups of students are engaged in different kinds of activities at a given time (individually composing a note, small group discussion and problem solving, whole class discussion). But at some point the teacher apparently leads open discussions and provides "benchmark lessons" (diSessa & Minstrell, in press; Minstrell, 1989), to synthesize the understanding achieved so far. She probably intervenes to keep the dialogue from wandering too far off track, for example, to clarify lingering misconceptions, or to suggest other data sources. However, the teacher's role is not described very fully in these papers.

(2) Learning at advanced stages raises questions concerning how much direction or structure would be needed if all members of the group have some or a great deal of expertise. For example, if the dialogue is among members of a scientific research group, a group member often assumes the role of synthesizing and organizing new understanding. In an adult training session, perhaps this role could be assigned to one of the members, or it could rotate among all members, such that no expert guidance is needed.

(3) Is it more productive of reflective thinking if students have unlimited time on the computer or if they are restricted to, say, 30 minutes a day, as in CSILE? Empirical research is needed on this issue. It is quite possible that more thoughtful and better articulated ideas are entered into the communal database if there are constraints on freedom to access the database.

(4) The compatibility of CSILE's computer supported dialogic model with Brown and Campione's classroom management strategies should be noted. Indeed, these two programs, have been combined and further extended to include the Jasper macro-contexts for mathematical problem solving, described below. These three programs form the core of an experimental

school demonstration project currently underway as part of the McDonnell Foundation's Cognitive Studies for Educational Practices program (see McGilly, 1994)

5.2 More Specialized systems

5.2.1 The Jasper Woodbury Problem Solving Series

Authors

Cognitive Technology Group at Vanderbilt

Reference

CTGV (1992), The Jasper experiment: An exploration of issues in learning and instructional design. Educational Technology Research and Development, 40, 65-80.

CTGV (1993). Anchored instruction and situated cognition revisted. Educational Technology, March, 52-69.

CTGV (1994) From visual word problems to learning communities: Changing conceptions of cognitive research. In K. McGilly (Ed.), Classroom lessons: Integrating cognitive theory and classroom practice. Cambridge, MA: MIT Press.

Grade/Age Range

Middle school students

Content domain

Arithmetic story problems. Jasper provides complex, video-based story problems anchored in a realistic situation. These "macro-contexts" provide a multi-level, instructional context within which students and teachers collaborate in decomposing the problem into do-able subproblems to construct a workable solution path. Unlike traditional, cryptic story problems, Jasper problems afford opportunities for problem exploration and problem generation from differing perspectives. Instead of a single correct solution path, there are several alternative wasy to tackle these problems - as is the case when dealing with complex problems in the real world.

Which educational or cognitive theory is it derived from?

This system is derived from constructivist theories, emphasizing opportunities for situated and collaborative learning. A key idea in Jasper is the notion that learners will be better able to reason mathematically when instruction and practice is "anchored" in a real-world problem-solving context. Jasper can be applied within various teaching models, from those emphasizing basic skills first through structured problem solving, to guided generation of problems and subproblems. However, Jasper's features will only be fully used in less directive, generative contexts.

Instructional context
Intended for whole classroom and small group problem-solving sessions, each problem takes at least two 1-hour class periods. Initially designed for 5th-6th grade math, it is now being extended upwards and downwards.

Where is the locus of control?
The locus of control is in the system which offers a video tape or disc presentation of a story; however, users can interrupt the presentation for discussion or search through it afterwards for specific data needed to solve the problem or subproblem. The degree of user control differs according to how the classroom is structured.

What does the system instruct ?
The system is designed to enhance conceptual and strategic knowledge: users need to identify and define problems, generate relevant subproblems, figure out relevant data for subgoals on their own. The goal is to help students become independent problem solvers, capable of figuring out the information that is needed and seeking it out on their own. Having students work in collaborative groups is supposed to aid their monitoring skills because they keep each other on track.

How is the content structured?
Pairs of related problems are presented together to foster transfer and analogous thinking. Content is not structured otherwise because the system is designed to foster open-ended problem-solving skills such as those one encounters in real-world settings. However, Jasper can also be embedded in basic skills curriculum to provide practice on previously learned component skills. Future work will extend the range of difficulty of problems: for example, using whole numbers instead of decimals, less complex, alternative solution paths, etc..

What does the learner construct or manipulate?
Learners set up a solution space, define the subgoals for complex, open-ended problems. This approach is assumed to teach better reasoning, incorporating broad use of calculators and computer-based calculating tools (spreadsheets, graphing programs), and enables extensions to analog problems or even other domains.

How is prior knowledge used?
Solving these open-ended problems taps students' general world knowledge and a broad range of math skills, which mostly have been learned previously, though some may need refreshing. Sometimes new skills are instructed in as needed in a particular context, which serves to make their usefulness obvious to students.

Opportunities for transfer
All of these are up to classroom teacher:

- analog problems may be presented;
- students may be asked to make up analogous problems;
- an "adventure maker" is currently being designed to allow students to create their own adventures under given constraints (e.g., flying an ultralight with a 4-gallon, instead of a 5-gallon tank);
- extension problems are suggested to other domains: for example,, to historical or current events; extension videos are also being developed.

Amount and type of support, when provided
Support is provided by the teacher, or by members of the collaborative group, not by the system, and may range from very structured (e.g., setting up the problem with blanks to be filled) to the teacher functioning as a guide, model, and co-learner. At the outset the teacher often needs to supply a prompting question to suggest alternative solution paths to try. Support may also take the form of structured exercises.

Range of tools that are available
Videodisc or videotape format. Videodiscs allow easy search, freeze-frame, random access. Since computers are not required, this is a low-tech, low-cost system which is easy to implement.

Forms and levels of interaction with tools
Users do not directly interact with the system beyond the usual VCR functions which allow users to access particular information (see above).

Errors and feedback
The authors assume that errors will be corrected by other members of the collaborative group or the class as a whole; they are treated as opportunities for learning. Feedback also occurs in discussion during which students are required to explain their solutions and justify their reasoning.

Interest, aesthetics
Lively, interesting story contexts; students seem very motivated and interested in the group problem-solving activity; all levels of ability get involved.

Assessment of learning
Preliminary findings from assessment studies currently in progress are reported in the 1994 chapter, which also provides references to more detailed reports of the results.

An ambitious plan for assessment is also outlined in that chapter, which involves comparisons of differences in instruction (Jasper vs.

traditional story-problems) in terms of various kinds of transfer ability (ability to solve a related Jasper problem, the extent to which understanding of general principles is shown in doing so, spontaneous application of knowledge acquired in working with Jasper to out-of-school problem situations. The basis for detailed assessment of transfer rests on a careful cognitive analysis of the solution structure of a Jasper problem, i.e., a specification of the plans, goals, and decisions needed to solve each episode.

(1) Baseline studies were conducted with college undergraduates and 6th-graders with high achievement scores in mathematics to provide an estimate of the efficacy of traditional school training as preparation for solving a complex Jasper problem. Assessment was based on personal interviews after viewing a Jasper problem video in which students were asked a series of problem questions. They were asked to think aloud as they attempted to answer the questions, and their protocols were scored, not only for solution success but also for whether and how they dealt with specific subgoals. The results showed important weaknesses in both subject goups in the ability to set up the problem space of a complex problem, namely the kinds of problem decomposition skills, as opposed to purely operational skills, that Jasper environments are designed to foster. Clearly, students need much more experience and practice with complex problem solving.

(2) Comparisons of instruction were conducted with high-achieving 5-th grade students.using either the Jasper video context or working on traditional word problems of the same problem type Jasper students scored much higher in tests of both near-transfer (analogous to the Jasper problem they had solved earlier) and a related problem with a different goal structure. However the two groups performed at similarly high levels on individual one- and two-step word problems. Jasper students also tended to find more optimal solution paths.

(3) A different kind of transfer was explored by exposing some of the Jasper students to "what-if" versions of the orignal problem they had worked on. These problems varied the conditions such that students had to figure out which aspects of their previous experience was relevant to the current problem. The results suggest that the what-if scenarios help students develop more flexible knowledge representations to support their mathematical reasoning.

(4) Teachers reported numerous incidents of students spontaneously relating Jasper experiences to activities in other classes or even out of school. However, a more formal follow-up study of students' tendency to make connections between what they had learned and an entirely different kind of situation (evaluate videos similar to the Jasper video) found that neither students with the Jasper training nor without it (i.e., who only viewed the video) tended to make connections without explicit prompting.

(5) Current research is investigating whether there is an increase in efficiency from learning in an anchored instruction context, such as Jasper, to unrelated domains.

General Evaluation

Although interaction with the system is minimal, the videos provide rich and challenging opportunities for group problem solving. However, the success of the Jasper system depends critically on teachers who are able to exploit its potential, as is the case with the Community of Learners model in both Brown and Campione (1994) and Scardamalia and Bereiter's (1994) implementations. The importance of teacher training became apparent as Jasper was implemented in a large number of schools and has led to intensive two-week training workshops for teachers. Implementation data, largely based on teacher reports, news articles, and a few classroom observations, are being collected. These anecdotal accounts are generally positive, and suggest that the system stimulates many spin-off activities (field trips, community projects). However, much more research and observations are needed on the issue of how to re-engineer traditional classroom instruction. In particular, we must consider what type of instruction (e.g., situated, guided or open exploration, or direct instruction) is best suited for what kinds of domains and learning goals (cf. Weinert, 1995).

Large-scale testing of the Jasper system with matched Jasper and non-Jasper-using classrooms are currently underway. Beginning, mid- and end-of-year testing employs both standardized tests and pencil and paper tests of complex problems solving skills developed by the researchers (e.g., tests of planning and subgoal comprehension), traditional word problems (from one to several steps), and assessments of students' attitudes toward mathematics. Preliminary results show significant gains on most measures. In general, however, the researchers found less evidence of flexible transfer than hoped for (p. 199).

The CTGV anchored instruction approach has been extended by holding competitions between classrooms working on Jasper problems in different schools. Students' presentations and justifications of their solutions are communicated through video-based teleconferencing facilities. The idea here is to develop larger communities of learners that allow students to test their ideas more broadly, and to build on their understanding by comparing their answers to those in classrooms throughout the community.

5.2.2 JANUS, a design environment for architectural design

Authors
Fischer, G., McCall, R., & Morch, A.

Reference
Fischer, G., McCall, R., & Morch, A. (1989). JANUS: Integrating hypertext with a knowledge-based design. In Hypertext '89 Proceedings, (pp. 105-117). Pittsburgh, PA: Association of Computing Machinery.

Content domain
This system is an example of the construction kit approach, which provides domain-level building blocks combined with a critic. The overall goal of this approach is to support solving of ill-designed problems in a wide variety of domains. Users can interact with the system, actually do programming without needing to learn programming language. They are dealing with familiar terms, and the system has the domain knowledge to correct design errors and to provide explanatory feedback (why a particular action is a poor idea, offer suggestions). The learning process is made more efficient and more effective because it takes place "on demand", while dealing with a particular task or subproblem. The learner does not have to have the knowledge beforehand nor to interrupt the problem solving to figure out how to access relevant knowledge. The content domain covered by JANUS - architectural design, specifically kitchen design - was developed as a demonstration project, but the ideas are broadly applicable to other, even fairly open-ended domains.

Which educational or cognitive theory is it derived from?
Meaning construction, guided discovery, situated learning, collaborative problem solving, learning on demand as information is needed - all characterize learning throughout the life span, according to Fischer et al. Therefore, critic tutors have broad potential uses, not only in school, but also in the workplace and at home.

Instructional context
Individual tutor or small group collaboration

Where is the locus of control?
JANUS is not an expert system guided by top-level goals and designs. Instead, the user controls the system at all times: the critic can be turned off by the user and the system used only to construct a design. Likewise, a good design can be chosen from the catalog which requires only minor modification, thus activating few suggestions or criticisms from the critic. This approach aims to combine construction and argumentation (Hypertext network) which is relevant to the current construction task. The argumentative component can

be explored in brief or at length by pursuing links to further information in order to further understanding of design issues.

What does the system instruct
(1) Conceptual knowledge building in a specific content domain with explanations which focus on underlying principles.

(2) Knowledge of system tools and resources is also instructed.

How is the content structured?
The designer can either construct from scratch or modify an existing design.

What does the learner construct or manipulate?
To construct a new design the designer chooses building blocks from a "palette" forming a design unit and moves them into the work area. To modify an existing design the user brings one into the work area from a catalog which contains both good and poor designs. Poor designs especially help learning the design principles because they will elicit a large number of criticisms.

How is prior knowledge used/enhanced?
Learners can pursue problems at different levels of difficulty, self determined, or explore the argumentative function as desired. They can better exploit the high functionality of the system without having to first master complex tools. The system is able to accommodate a range of ability and knowledge.

Opportunities for transfer
Presumably transfer will be enhanced by understanding of underlying principles and by learning in a variety of contexts.

Amount and type of support, when provided
Support is in the form of feedback, both positive and negative, provided when the critic system detects a design problem. The critic acts like a coach, offering suggestions and explanation of principles underlying a design decision. Users can choose increasingly complex design problems, according to their own goals and fade out use of the critic:

- Easy designs to get started: there is no need for prerequisite skills, such as system building.
- Challenging but attainable goals: one can work in richer environments without interrupting the problem solving or construction to search for information. Learners can gradually relinquish dependence on supports because they are also learning how to locate and use the relevant information resources.

- Scaffolding, not only in graduated difficulty of the problems, but also in the complexity of concepts: not too much new information is presented at one time; abstract concepts are presented in a concrete context.
- Individual differences in background knowledge are also accommodated.

Range of tools that are available
The goal of all LOGO based systems is to make available the whole richness and power of the system, with minimal programming knowledge. The system encorporates the idea of "no threshold, no ceiling", allowing learners to advance through increasingly complex microworlds (Burton, Brown, & Fischer, 1984), acquiring both programming skills as well as domain knowledge .

Forms and levels of interaction with tools
Construction kit - users may build a design, such as a kitchen, on the screen from component units from scratch or modify an existing design chosen from the menu.

Errors and feedback
Errors are opportunities for learning. Feedback from the computer critic supports the learning process (see above).

Interest, aesthetics
Little information is provided, except to note that in informal assessment users found the system easy and fun to work with.

Assessment of learning
No formal assessment study has been reported.

General evaluation
The AAAI paper provides only a very general description of use and evaluation by amateur and professional designers. No specification is included of how their learning was assessed, though probably an opinion questionnaire was used. The following limitations have been noted by the system builders:

(1) JANUS only allows viewing at the room level, not at higher or lower levels, so one cannot zoom in or out.

(2) It does not allow simulating the use of the design, a technique often used by designers.

(3) It does not allow designers to input their own goals, preferences, specifications for a particular problem, for example, to manipulate competing goals such as cost and feasibility tradeoffs, thus it remains limited to abstract

Abstract

Principles of Learning in Multimedia Educational Systems

The principles of learning summarized in this document are derived from theories about comprehension and learning that have guided research in cognitive science in recent years. With this summary we seek to make these theoretical ideas and techniques from cognitive research more easily accessible to people who are engaged in designing educational software programs and other kinds of multimedia learning tools.

Our assumptions about learning are based on the constructivist perspective that learning is an active process by which learners construct their own meaning from a particular situation or from various media, such as texts, lectures, pictures, video, graphs, and so on. Mental activities that promote active learning and deep conceptual understanding depend on several, complexly interacting factors: what kinds of processing strategies the learner employs, the form in which the learning materials are presented, and the social-instructional context in which learning takes place.

We then describe a few examples of instructional programs that promote active, engaged learning. These range from general classroom management systems for collaborative learning to more specialized systems, including individual tutors.

Finally we discuss the problem of evaluating educational software systems, which in our opinion remains a major weakness in this field. We provide some ideas about how learning can be assessed derived from psychological research and conclude with a brief discussion of the value of linking educational software design and evaluation with theoretical notions about how learning takes place.

situations. Note, however, that these criticisms are being addressed in extensions of the system.

(4) The system does depend on a diagnostic system, though not as extensive as student modeling. Nevertheless, a huge amount of knowledge may hve to be built in, depending on the open-endedness of the domain.

(5) Knowing when to interrupt with a critique, how often, what feedback is needed are also a very complex issues. For example, sometimes it may be better to let users discover and remedy errors on their own. Can users also be expected to know how much information to seek? And do they actually explore issues in depth of just far enough to get the job done?

(6) More exploratory research is needed to determine for what kinds of domains and what level of learning this system is best suited. Certainly it is a powerful tool for instructing a well-defined set of rules and priniciples in domains, such as environmental design, or subdomains thereof. It is probably not optimal for acquiring basic principles in broad, open-ended domains because of the amount of knowledge that must be built into the critic. In addition, evaluation studies are needed to specify what is learned, how well, and how cost effectively.

5.3 Hypermedia/HyperText

Hypertext differs from linear text in that it provides many choice points from which the reader can branch out according to his or her interests or special needs. For such a system to be effective, good maps and navigational devices must be available. The advantage of hypertext lies in this tailorability to individual goals and requirements. There are some dangers, however. Hypertext allows readers to design their own text, to organize their own book. While this has some obvious attractions, organizing a text is no easy task, as any author can tell, and to put this burden on an inexpert reader is not without its risks. In many situations, the author knows best how to organize his or her subject matter, and most readers can do no better than follow this organization, even when they have the opportunity to deviate from it (Foltz, 1992). In the worst case, the reader will become confused and lost in the system. Moreover, empirical evaluations of hypertext systems have for the most part not shown them to be superior to linear text (for a review, see Foltz, 1992). However, there are exceptions. For certain tasks, a well-structured hypertext system can be much superior to a conventional book. Most notably, Superbook, a hypertext browser which is based on cognitive research and has been extensively and iteratively evaluated, has proven to be highly effective for locating dynamically needed information for problem solving tasks in various domains, such as chemistry and statistics (Landauer et al., 1993).

5.3.1 Cognitive flexibility and hypertext

Author
R. J. Spiro

Reference

Spiro, R. J., Vispoel, W. P., Schmitz, J. G., Samarapungavan, A., & Boerger, A. E. (1987). Knowledge acquisition for application: Cognitive flexibility and transfer in complex content domains. In B. K. Britton & S. M. Glynn (Eds.), Executive control processes in reading. Hillsdale, N. J: Erlbaum (pp. 177-199).

Spiro, R. J., & Jehng, J. C. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix and R. J. Spiro (Eds.), Cognition, education, and multimedia: Exploring ideas in high technology. Hillsdale, NJ: Erlbaum (pp. 163-205).

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1991). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill structured domains. Educational Technology, 31, 24-33.

Content domain
Hypertext, or random access presentation of text, is particularly suited for concept learning at advanced stages of knowledge acquisition when the learner is faced with more complex and ill-structured content, and applications to specific cases or examples are less straight forward.

Which educational or cognitive theory is it derived from?
According to the cognitive flexibility hypothesis, flexibility in the knowledge base is necessary in order to apply complex knowledge to specific cases in ill-structured, open-ended domains (e.g., medical diagnosis), since knowledge structures will have to be reassembled in various manners depending on the dictates of the specific cases. These ideas are a version of constructivism which emphasizes the situation specific assembly and reassembly of components of prior knowledge. Accordingly, the knowledge acquisition process should support maximal interconnectedness of knowledge representations, whereas traditional modes of representation and learning tend to oversimplify and compartmentalize knowledge. Concepts with multiple connections in the knowledge base form a web-like network, or "criss-cross conceptual landscape" which is best supported by nonlinear presentation of subject matter such as that afforded by hypertext.

Instruction context
Individual tutor

Where is the locus of control?
Users choose their own path through the subject matter by accessing whatever topics they wish. Hence, the learning process is undirected.

What does the system instruct?
The focus is on forming linkages among concepts.

How is the content structured?
The content is organized into cognitively manageable chunks of information without oversimplifying the content. The implementation developed by Spiro et al. was designed to teach an advanced understanding of the film, *Citizen Kane*. The system is organized around small segments, or minicases, from the film and 10 critical themes relevant to understanding each minicase. The system provides the student multiple perspectives with which to examine the minicases or use of a particular theme. For example, students can examine the various themes relevant to a minicase or search the minicases by theme or combinations of a theme, thus combining and recombining elements of the film according to their own particular purpose.

What does the learner construct or manipulate?
Learners thus structure their own conceptual understanding rather than learning information in a rigid, pre-organized form. They choose how to link elements of presented information, which, according to cognitive flexibility theory, should result in the construction of a knowledge base composed of multiple interconnecting links to other concepts, for example, cross references to other themes.

How is prior knowledge used?
Given that learners design their own learning path, this system draws heavily on prior knowledge, interest, and personal goals.

Opportunities for transfer
Spiro et al. claim that broader experience during the learning phase, for instance, by traversing through multiple perspectives on the subject matter, should facilitate applying the knowledge to novel cases. That is, multidimensional analysis supports multiply linked conceptual understanding and avoids rigid compartmentalization of knowledge. Thus the learner develops more efficient access to relevant knowledge in memory. In using this hypertext presentation learners are constantly required to apply new information to new cases during the learning phase.

Amount and type of support, when provided
Very little guidance is provided by this hypertext system. The only structure is in terms of how the content is decomposed.

Forms and levels of interaction with tools
Hypertext simply provides non-linear, random access to the material. The user chooses which segments of information to read in which order by pressing on topic buttons. Apparently there are no labeled links, instead they are implicitly formed by the student.

Errors and feedback
There are no opportunities for error correction or feedback within the system.

Assessment of learning
Assessment by Spiro, Vispoel et al. (1987) consisted in comparing a "criss-crossed" hypertext group of students with a control condition in which information was presented in a well-structured, linear textbook fashion. The hypertext group outperformed the control group on all six transfer measures which required applying the knowledge to new kinds of problems. However, the control group was better on tests of factual memory. The researchers concluded that their knowledge was structured too inflexibly to be easily accessed for novel applications.

General evaluation
The nonlinear presentation of material via hypertext potentially does encourage more constructive, inferential kinds of processing, due to the necessity of figuring out the relationships between segments of information. It is of course true that these are the kinds of processes required for representing knowledge at the level of the situation model, as opposed to a more superficial textbased understanding. However, the lack of guidance can be problematic for all but the most advanced and self motivated learners: What is to prevent learners from constructing the wrong kinds of links between concepts, that is, to construct an inaccurate situation model? And what is to prevent students from reading the information in a thoughtless manner, without really thinking about the ways concepts, or minicases and themes, could be linked? Thus, passive reading is just as possible in an unguided hypertext system as with linear text, and the consequences of unstructured content may be much more damaging.

In a recent study by Shapiro (submitted), the beneficial effects of a nonlinear text presentation did not extend to tests of learning involving inference generation, even though the hypertext groups had developed more associations between concepts than the linear text group. Clearly, a good deal more and very careful research is required to determine whether, under what conditions, and for whom hypertext presentation may result in better understanding and learning. In particular, differential effects of readers' background knowledge should be considered, since high-knowledge individuals may well respond differently to the lack of structure than low-knowledge readers, as was the case for coherent and less coherent text (see McNamara et al., in press, and other studies discussed in section 4.1.4.3).

6. Evaluating Interactive Media for Education

6.1. Defining the context of evaluation

Evaluation of educational methods and tools is, as Reeves (1990) remarks, inherently dependent on historical, political, and situational circumstances. An evaluation of the use of a particular medium, say the radio research of the 1950s or the television research of the 1960's (e.g., Hovland, Lumsdaine, & Sheffield, 1949; Schramm, 1977), had to justify the investment in the medium to educational administrators. This meant several things: the evaluation had to be expressed in a language meaningful to these decision makers; evaluations were based on conceiving teaching and learning as disseminating and taking-in of basic facts and knowledge, and the medium had to be proven more efficient, or at least more cost-efficient than traditional classroom lessons. There was a strong assumption that general knowledge, such as language skills, historical knowledge, math skills, and basic science facts had and could be internalized by individual learners to prepare them for the application of these skills and knowledge in the professional world. There also was (and still is) an assumption that knowledge of this type could be measured in standardized, general tests.

The following criticism of these studies, put forward by Clark (1983), implies these assumptions, and points to problems with this type of research:

"The best current evidence is that media are mere vehicles that deliver instruction but do not influence student achievement any more than the truck that delivers our groceries causes changes in our nutrition. ... It seems reasonable to recommend, therefore, that researchers refrain from producing additional studies exploring the relationship between media and learning unless a novel theory is suggested."

Clark was not the only reviewer to lament the inconclusiveness of evaluation studies based on the paradigm of comparing a medium of interest with traditional classroom lessons. He is joined by others (e.g. Cronbach, 1975; Hoban, 1958; and Reeves, 1986, 1990). As sources for the inconclusiveness of the studies, confounds, such as teaching method and novelty effects, and inappropriate manipulations and measurements are discussed. However, the goal of evaluation to convince decision makers to invest in a medium as a superior method of knowledge dissemination did not receive criticism, as it defined the proper political context for these studies up until the eighties, and even in some senses until today.

In the last twenty years, the cultural and political context for which the educational system should prepare has changed dramatically. A large and growing proportion of people now live and work in electronically mediated environments. Job-relevant information abounds, and changes constantly. Information can be accessed instantly even from remote locations, and it can be shared between co-workers in geographically distant locations. Work in the information age does not rely as much on the retrieval and application of general or specialized sets of knowledge from the minds of individuals. Instead, precise and specific information can now be retrieved electronically by anybody having access to the resources, or it is embodied in tools, such as CAD tools, or even in simple functions, such as spell checkers.

This leads increasingly to the elimination of jobs that ask for the application of rote skills and knowledge in paper and pencil environments, where product cycles were slow, and the need for sharing of the results infrequent. Today, many tasks can be completed instantly and automatically so that the main work emphasis is at the level of communicating the results, and negotiating decisions based on it. Thus, collaborative use of external knowledge repositories, for joint construction of problem solutions is and will be central to the work life. Industry leaders are increasingly emphasizing this shift in job skills, namely, a need for computer literacy, and the ability to collaborate, and to communicate successfully with the assistance of electronic media (e.g. Time Magazine, Spring 1995).

This situation defines part of the socio-cultural context for developing and evaluating educational media and methods today. A second educational need is remedial. Especially the public educational system has been challenged recently with the task of providing adequate basic education to all students. It seems that the 'traditional classroom lecture' does not meet the changing needs of a large proportion of students, especially students whose basic educational needs are not met by their families. Another goal of educational change therefore is to provide learning environments and methods that will integrate the activities of a diverse student body and engage all student in meaningful and knowledge constructing tasks.

6.2 Three goals of evaluation

6.2.1 Measuring the desired effects

How does this context define the goal for evaluation of teaching methods and media in the classroom? At first, researchers and educators will have to continue proving that electronic media, computers and access to networks are desired, or cost effective teaching tools. Even though computers, and even networks have been introduced into a large number of classrooms (e.g. Time Magazine, Spring 1995), public educators continue to struggle justifying updating of quickly outdated equipment (e.g. Bulkeley, 1995). But, the question is also how to embed computers usefully in the classroom, how the computer will change the classroom interactions, and how these anticipated uses will lead to the type of skills that industry expects their new generation of employees to bring into the workplace. Furthermore, there are many systems and services that can be installed on computers and networks, and there is a need to decide which ones will influence classroom interactions in a desired way. There is also a need to show how the use of such systems influences students achievement differentially, and whether there is a chance to increase the participation and achievement of low-achieving students.

6.2.2 Describing the patterns of use

Another aspect of evaluation concerns itself with the use-aspect of systems. In other words, to design multimedia systems successfully, designers as well as educators will need to know how a system will be used in a certain situation. This question is especially important for systems designed from a constructivist perspective, because many of the systems (see Section 5) above, don't force a student through a set of drills, but enable the student to create documents, programs, or find their own way through a documentation (e.g., Adams & diSessa, 1991; Bruce & Rubin, 1993; Shapiro, submitted). On the other hand, systems are not necessarily designed without intended uses. For example, in the Electronic Quills writing package, designed by Bruce and Rubin (1993, p. 45), educational goals (such as learning to plan) were implicit in the provision of software features (such as a planner, that helps in structuring a paper) It is necessary to know whether these features are used at all, in which situations, by whom etc. For the designers of the software it is also important to know whether there are usability hurdles to the use of any functions.

6.2.3 Relating patterns of use with effects

Finally, to learn about the design of interactive educational environments in an optimal way, system designers as well educators need to know which particular uses of a system by which users and in which environment leads to which particular result. An accumulation of this type of information enriches our knowledge about known methods of system design and use to achieve specified learning outcomes. As of now, there are still few empirical studies that address this causal connection in a satisfying way (but see Reeves, 1993).

6.3 Methods of evaluating learning

In this section a range of measures and methods to achieve the three goals of evaluation will be suggested. The list included here is not and cannot be exhaustive, as many of the variables and behaviors that can be measured and observed depend on the particular system under investigation. What we are trying to convey is the idea that for a study to be conclusive, researchers will have to be both creative, and analytical. They will have to analyze the possible uses and effects in an appropriately detailed way, and they have to be creative in inventing measures that tap those variables appropriately. The discussion here is not meant to provide a complete cookbook, but it should provide some examples and issues to get started.

The lists provided below are organized along the following lines: A general method or measurement is listed and explained briefly, including the reason for including such a measurement. When available, a brief example from the literature will be given.

6.3.1 Measurement of effects of using the system

6.3.1.1 Standardized tests and measures (e.g., CTBS, SAT, ACT scores).

What is it?
Standardized tests, such as math and verbal SAT scores and various reading comprehension instruments are designed to test skill level in a generalized skill or knowledge area. Tests like these are used nationally, and they are usually designed to avoid local or other biases that are other than skill related. Standardized tests usually come with interpretations for specific scores, that is, there are guidelines that rate performance compared to a national level.

Why use it?
Standardized test do not have to be designed from scratch and are easily administered. They are designed to be administered repeatedly without inducing test-related learning effects (several versions of a test). Educational or grant agencies may request a measure that allows comparison with a known standard. It may serve as a good baseline measure of performance, and it may be interesting to examine differences between outcomes using these general measures and more specific tests, designed to measure performance of particular knowledge or skills (see below).

Problems
These tests are general purpose instruments. They may not test specific learning results. They are fact oriented, not process oriented, and therefore they may not capture changes in motivation, approach to learning and knowledge oriented tasks. They have been criticized for being culturally biased. The may be appropriate for selection based on comparison with a standard, but they do not necessarily allow for meaningful interpretation of the acquired knowledge or skills.

Recommendation
Use for documentation and backup purposes. Use to compare achievement of treatment vs. control performance. Use to track change in performance over time. Supplement with more specific tests that measure other more interesting aspects of learning.

Example
Scardamalia et al. (1994, see Section 5.1.2) used the composite math and language skills scores of the Canadian Test of Basic Skills (CTBS) to demonstrate that CSILE students were performing at least as well as students in a more traditional classroom. Results showed that CSILE students had an advantage with respect to verbal skills, but not with respect to math skills, as expected. Explanations for these differences were gained from other, more specific measurements (see below).

6.3.1.2 Simple factual knowledge

6.3.1.2.1 Recall and recognition.

What is it?
Recall and recognition are the classical tests of memory for declarative information learned from texts. (e.g., van Dijk & Kintsch, 1983; Kintsch & Welsch, 1991). In recall, subjects are asked to write down or verbally recall as much as they can from the studied material. Recall measures can be used to estimate how much information has been retained from a set of material. Various coding schemes can be used to classify the recalled information with regard to its position in the text structure. In recognition, subjects have to respond to specially crafted items, and verify that this is a true or a false statement, according to previously studied material. Recognition tests also measure memory for declarative information. For this measure, recognition items need to be constructed. Recognition items can correspond to the studied material in various respects: they can be verbal items, paraphrases, or sentences that contain true inferences, as well as items representing false information Recognition rates associated with various classes of items can be used to estimate the retained knowledge on different levels of representation: surface information, propositional representation, or a representation of a model of the situation described in the material (for a detailed explanation of the measure, see Kintsch & Welsch, 1991).

Why use it?
These are important measures to use when the goal of an instructional system is the construction of declarative knowledge structures. Items can be designed to test for different levels of learning, especially, whether students are able to actively use the presented information for decision making or inferencing. Many hypertext or multimedia applications have exactly this goal in mind, and recall and recognition methods may be the appropriate measures for these situations.

Problems
Scoring of recall protocols and the crafting of recognition items may be very time intensive, depending on the length of the recall protocols and the number of classifications used. Recall may only measure the surface

representation of what was learned, and not necessarily reflect how the knowledge was integrated into an overall knowledge structure. Both, recall and recognition only measure the net result of a learning episode. They do not directly lend themselves to interpretation of the meta-strategies or learning methods used, and they do show how and whether the acquired knowledge can be used flexibly in problem solving (e.g., Mannes & Kintsch, 1987).

Recommendations
Looking at recall and recognition scores alone might not be very interesting. It is very important to compare across different presentation versions of a system, or to correlate recall and recognition rates for various levels of information with variables that describe the mode of presentation (e.g., whether it was hyper-linked, presented in text format, presented in graph format, presented as a simulation, occurred in the highest level of the hypertext structure, etc.). It might be interesting to use these measures at various forgetting intervals because the advantages of some presentation methods might lie in better long-term retention of a certain type of knowledge.

6.3.1.2.2 Measures of conceptual structure

What is it?
In these tasks, subjects are usually asked to respond to items containing single concepts, such a words, phrases, or pictures. In one type of task, cued association, subjects are shown a list of concepts (usually from the material studied) and asked to say the first association that comes to mind (e.g. Ferstl & Kintsch, 1991). If the same list of items (in different orders) is presented before and after studying a set of materials, differences in the structure lead to conclusions about a change in the conceptual structure that makes up an individual's declarative knowledge in that domain. Similarly, subjects can be given a set of cards corresponding to single concepts, and asked to sort them into meaningful groups. Changes in groupings can again be used to determine changes in conceptual structure (Britton & Gulgoz, 1991). Several statistical methods exist to evaluate these changes numerically.

Why use it?
These measures are especially interesting in situations where the goal of an instructional method is to change a previous conceptual structure, or to show that a method leads to a good integration of new knowledge into an old structure, because an assessment of structures can be made before and after studying of material. This type of result could not be gained from recognition or recall measures for several reasons, such as test-learning from repeated exposure to the material, and the fact that a text can not be recalled before it was read, even though the concepts in the text may be known in some way.

1. Overview

In this paper we review the literature on learning and instruction in order to provide designers of multimedia instructional systems with guidelines for optimizing learning. We try to show not only what has worked and what has not worked in the past, but to offer a theoretical framework within which these results can be interpreted and which can serve as a guide for future work. How people learn provides a set of constraints for the design of any instructional system that is at least as important as other considerations that play a role in the design of multimedia instructional systems. Our goal is to offer a guide to the practitioner as to the nature of these constraints. This requires that the practitioner become at least somewhat familiar with learning theory and the psychology of higher-order cognitive processes. With this goal in mind, we discuss and selectively review the relevant literature. Thus, this paper is not a comprehensive treatise on learning and instruction, but rather is oriented towards the needs of systems designers who must deal with instructional issues.

We take learning to be the formation and modification of mental representations. In particular, we are concerned here with abstract, symbolic mental representations that play a role in much school learning and many job training programs in industry. The medium of learning may be classroom instruction, a text of some kind, or a multimedia system combining the written and spoken words with graphs and pictures, still or video. The presentation in both texts and multimedia systems is typically linear, as in a movie or book. However, modern computer systems also permit nonlinear, learner controlled presentation sequences, as in hypertext (nonlinear text) or hypermedia (nonlinear multimedia systems). The principles of learning which we propose and discuss in this article are quite general and are not restricted to multimedia systems. However, we make an effort to point out their special relevance and role in design of multimedia instructional systems.

1.1 Forms of Learning

Even the simplest organisms are capable of learning. Needless to say, they do not learn the same way as people do. People are complicated, however. They can learn in the simplistic manner of paramecia, but they also can learn in many other ways, like chickens, like chimps, and most importantly, in ways peculiarly human. Thus, any principles of learning must be prepared in multiple editions: how people learn by habituation, by

Example
A very creative and informative use of a cued association test was made in a study reported by Shapiro (submitted). In hypertext research the claim is often made that presenting material in a flexibly linked web of conceptual nodes with links between them, will impart a similar knowledge structure to the students that study in this type of environment. Shapiro's project was concerned with supporting this claim by comparing learning from different hypertext systems that implemented different links between concepts. After studying the domain in this way, her subjects answered to a cued association task, where they were given topic names from the system and had to answer with associations that immediately came to mind. The given associations were than categorized by whether they were contained in textual links only (where the two concepts were included in the same sentence), or whether the link was also implemented in a hyper link. The results from this analysis showed that concepts linked in the text AND linked through a hyper link had a much higher probability of being recalled than concepts that had only been mentioned together in the text. Comparing a hypertext version that implemented a typical hierarchical text structure with more flexibly linked presentations of the material led to more associations being formed in the flexibly linked versions than in the hierarchically linked version. This use of cued association avoided the difficulties of interpreting the implied conceptual structures, but showed that particular structures of hypertext can indeed influence the construction of knowledge in a characteristic and positive way.

6.3.1.3 <u>Application of factual knowledge:</u>
- inferences
- problem solving tasks
- analogies

What is it?
These measures ask students to draw on their newly acquired declarative knowledge for solving a problem in the domain, making inferences, or using it in analogies Inference sentences can be constructed similar to recognition sentences (see 6.3.1.2.1). Alternatively, carefully crafted questions can be asked that elicit whether students have actually made a connection between two assertions given in a text. Problem solving tasks usually ask students to provide a solution to a problem described to them. For example, if students study some material on causes of the greenhouse effect, then a problem solving task would ask them to solve a particular problem in this domain that was not directly discussed in the material, but that should be solvable on the background knowledge provided in it. Analogies, are a specific case of problem solving in which a general principle learned in the context of one domain can be applied in another domain to solve a problem. Flexible and appropriate use of analogies has proven to be especially difficult for students (e.g., Ross, 1989; Gick & Holyoak, 1983). Spontaneous use of an appropriate

analog structure is therefore a very good indicator of the construction of abstract and flexible memory structures.

Why use it?

These measures address exactly the type of flexible knowledge use that hypermedia systems and collaborative learning environments claim to support. Rather than just testing whether knowledge has been stored, performance on these measures lets us know how deeply it has been processed. Deeply processed knowledge should be well integrated into an existing knowledge base, that is, it should be accessible using many different cues, and usable in many contexts. Performance on inference, and problem solving tasks shows to what degree this goal has been achieved.

Problems

Constructing sensible inference and problem solving material is difficult and requires a good sense of what can be asked from a set of students. It is easy to make the questions too difficult and produce a floor effect. Material should be pre-tested and their construction can be time consuming.

Recommendations

It is important to work closely with teachers or educators in constructing good and sensitive test material. When unsure, testing materials should be pretested with small pilot samples. It may also be interesting to collect on-line verbal protocols from students trying to solve these problems. This might provide direct evidence for the types of knowledge that they are using in solving it. Finally, for a similar reason, it is useful to combine knowledge application tests with factual knowledge tests (see above). This enables researchers to see whether knowledge simply cannot be recalled, or whether it just cannot be used outside of the direct learning context.

Example

A paper by Brown and Campione (1994) discusses the use of problem solving and the use of analogy in evaluating a collaborative learning technique, called reciprocal teaching (see 4.1.2.2). Problem solving questions, for example, were to design a novel animal suited to a given environment (application of knowledge of evolutionary principles). The answers were scored according to taught principles: application of taught principles, generalization of principles to invent new similar ones, etc. Teachers and researchers also analyzed the student's processes while processing text. After reading portions of an expository text, students were asked to give answers to simple fact, inferential and analogy question; furthermore the type of explanations given to these answers was scored The results showed that students did not improve on comprehension of simple facts, but that they did improve in their ability to use that knowledge flexibly to draw inferences and analogies.

In a very different study, Gräsel, Mandl, Fischer and Gärtner (1994) evaluated the use of an interactive computer-based learning environment, THYROIDEA, that teaches medical students the knowledge needed to diagnose a dysfunction of the thyroid glands. As students are presented with more and more information about a case, they are asked to answer problem-related questions at different points during the presentation of the case. A 'coach' checks these answers and gives feedback about their accuracy. Students also have the opportunity to consult an expert segment about additional information on the symptoms. At the end of the program, students are asked for a final diagnosis of the case. In addition to other measures, these authors used the quality of the final diagnosis as an application of the factual knowledge to a problem solving situation as one criterion to evaluate the performance of the system. One interesting result of this study was that the use of the feedback to correct false answers was positively correlated with a good diagnosis at the end of the study.

6.3.1.4 Simple procedural knowledge: Problem solving. Testing for this type of knowledge assumes a different type of knowledge representation. All previous tasks assumed that the knowledge to be internalized or accessed was of declarative format, that is, some types of concepts, interpretations of concepts, or higher-level reasoning strategies. However, many learning curricula focus on the learning of procedural knowledge, that is, on knowledge of how to do a certain class of tasks Even though learning theories generally assume that procedural knowledge is usually acquired through some type of declarative representation first (e.g., Anderson, 1987; Fitts, 1964), the desired product is a quasi-automatic application of specific procedures that allow for fast and error-free application of a skill. Examples of such problem solving strategies are mathematical procedures (e.g., subtraction - Van Lehn, 1988; geometry - Koedinger & Anderson, 1990; low-level programming and debugging skills - Pennington, Nicolich, & Rahm, 1995). Procedures as such are usually acquired by repeated practice of a battery of identified subskills, feedback after an error has been detected, and then finally application of the subskills in more general task environments. These are the types of skills that Intelligent Tutoring Systems develop and support (e.g., Koedinger & Anderson, 1990). Even though our report aims at describing systems that support the acquisition and flexible use of declarative knowledge, a brief section on testing procedural knowledge is included for completeness.

What is it?

Problem-solving tests usually ask for the solving of problems in a relatively narrow domain that the students were especially trained in. For example, students might have just completed a training procedure on simple addition. In general testing then involves a battery of addition problems within the range of difficulty that was trained for. Measurements include the error rate, the speed of solving the problems, analysis of problem solving traces, and analysis of typical 'bugs' within those problem solving traces. Intelligent tutoring systems usually maintain a student model that traces performance on all these criteria. Students are only advanced to the next level of practice when they have reached satisfactory performance in the next lower skill level. Another method that is very popular is having students 'think out loud' (Ericsson and Simon, 1984) while problem solving. Taking verbal protocols provides access to the students problem-solving and reasoning strategies, especially when the to be acquired skill is still in the declarative phase of development, and therefore it is possible to decide why a problem could not be solved (missing information, buggy problem solving strategy, memory failures etc.).

Why use it?
Simple speed and accuracy measures can lead to conclusions about whether a skill has actually been acquired, and what stage of acquisition it is in. Verbal protocols, analysis of problem-solving traces, etc. lead to a diagnosis of where skill deficiencies occur, and can be used directly to enhance the underlying student model, or the curriculum.

Problems
These measures indicate whether the method or model teaches what it promises. They do not test whether the trained skill can actually be applied in a situation external to the tutor or class setting, that is, whether students will be able to recognize contexts in which the trained skills are applicable.

Recommendations
These types of tests are a must to assure that a minimum of skill acquisition has been achieved. Certain basic skills (such as math skills, literacy, basic programming concepts and procedures) are and will be necessary knowledge in the future. However, acquired skills should also be tested 'out of context', applied to problems that require transfer to make sure that students have acquired something useful in real-life situations 'Out-of-context' and transfer tests are discussed below.

6.3.1.5 Application of procedural knowledge to 'new problems'.

What is it?
Transfer problems can be designed at varying levels of difficulty, where each level determines a higher level of generality of the skill and flexibility in using it. These types of problem solving tests ask to apply procedural knowledge acquired in one type of task to a new problem domain. Let's say procedural knowledge about multi-column addition has been learned in the context of one type of application domain, and is tested in a very dissimilar application domain. A far transfer test would be to present a problem one of whose subproblems affords an addition procedure without explicitly asking for it. The addition training would definitely prove to be successful if students

could spontaneously recognize the type of problem and apply their knowledge without the affordance characteristics of the 'word problem', or 'tutor lesson' context. Another type of transfer problem asks students to apply particular concepts embedded in a skill to a different task. For example, one could ask students who have been trained to write program code to now debug somebody else's code, or one could ask students that have become proficient in multi-column addition for a solution of a multi-column subtraction problem.

Why use it?

If the educational goal is to impart basic skills that are widely applicable and can be flexibly used, then one should make sure to include situations which test this in the evaluation. Students may be able to produce answers to satisfy a teacher or tutor if the context is given, but they might not be able to export that skill to another situation. Similarly, if a training procedure has indeed involved some 'deep thinking' about a skill, rather than just rote practice, students might be able to solve a similar problem that draws on the same knowledge, but not on the same procedures (Pennington & Nicholich, 1991).

Problems

True transfer tasks may be difficult to design. It may be difficult to set up a situation so that the context cannot be used as a cue for the expected response. It may seem too cumbersome to take students out of the learning context (classroom, computer based instruction, word problems etc.) into another contex, to test how these skills can be applied.

Recommendations

If the evaluation is done as part of a class curriculum, it may be possible to measure some curriculum activity that emphasizes a different educational content (e.g., a history session, or a biology project), but can be used to measure a skill learned in a different class or different subject area. This would set a different context (content of class, teacher). This affords close work with the teachers in a school where the testing of a system is done.

Example

A good starting point for 'out-of-context' testing would be the design of the Jasper series (The Cognition and Technology Group at Vanderbilt, 1992; 1994, see description of the research program above in Section 5.2.1.). Jasper problem-solving contexts involve real-life problems, which are introduced in an adventure story video. Students are invited to identify with the protagonists' goals and solve a problem for them that involves many subproblems of differing contents. Even though the Jasper videos are designed as a type of instructional context, they can be used as a testing ground instead. Indeed part of the evaluation of the Jasper project (CTGV, 1994) involves tests of transfer to related or even quite different problems than the original training problems.

Pennington and Nicholich (1991) provide a good example of using a 'reverse skill' as a transfer test after training on an intelligent tutor. Here, the performance on the transfer test does not serve to evaluate the tutor, but to test a particular learning theory. However, performance on reverse skill could easily be used to test for differences in training on two different types of tutors or training methods.

6.3.1.6 Higher level knowledge (strategic or meta knowledge):
- planning strategies
- metacognition/self-monitoring strategies
- advanced comprehension strategies

What is it?
These are the domain independent general problem-solving strategies that have been very difficult to induce and test in past research (e.g. Singley & Anderson, 1989). However, many constructivist theories have the acquisition of such knowledge as their goal, which means they should at least attempt to show that methods suggested by them are approaching these goals. Planning strategies are general approaches to a problem, namely, recognizing that there are several subproblems to solve, and determining the interdependencies between the solutions to these subparts. They could also include the skill of determining that different people or groups of people should work on the subparts and report to each other for solving the overall problem. Self-monitoring strategies or advanced comprehension strategies help students to determine that they have reached an impasse either in comprehension or problem solving. If such an impasse is detected, expert problem solvers have analytic skills to determine what the current impasse is resulting from. It could be the lack of a piece of information, it could be the lack of a skill, or ambiguity in the information provided. One good way to monitor the application of these high-level problem solving strategies is to observe students' behavior, elicit verbal protocols, or even prompt them for explanations of a particular problem solution. Test problems or texts should have known problems in their representation, so that strategies to fix these problems can be expected at certain points in the problem solving procedure.

Why use it?
As mentioned above, constructivist theories claim to support the acquisition of such higher level metacognitive learning strategies, thus it is important to test whether they are really approaching this goal.

Recommendations
To elicit higher level problem solving strategies, there should be a clear definition of the strategy to be elicited. For example, a test for metacognitive strategies involving text comprehension would involve a text with known comprehension difficulties, or a test for good planning strategies would

involve the solving of a complex problem with many subgoals and interdependencies between subsolutions, such as the Jasper video-based problems (CTGV, 1994). Strategies can be elicited by verbal protocols, or by having several students work on a problem together. Their conversation about the problem can then be analyzed.

Example
To test the effect of the reciprocal teaching strategy, Brown and Campione (1994) gave their students difficult expository texts outside of their area of study and asked them for facts, inferences, gist, etc. at various points during the reading of the texts. The students' responses to these questions showed that students who had been using the reciprocal teaching strategy, improved especially on the inference, gist, and analogy questions.

6.3.3 Linking use of a multimedia system and effects on learning

6.3.3.1 Exploratory studies. If the study questions are open-ended, that is, if there are no clear predictions about how a system might be used and which of the functions will lead to some type of learning result, an exploratory research strategy is recommended. In an exploratory trial, a wide range of information should be collected to allow for discovery of unexpected variables.

With respect to usage patterns, video-taping all interactions with the system rather than observation only is a very good start. Having a video tape of the interactions available makes it possible to identify interesting behavior patterns in a first step of analysis. Once several interesting types of behavior have been categorized the tapes can be used to collect more details of these behaviors (frequencies, timing, etc.). If data are collected through observation of ongoing behavior (without recording), it is often difficult to collect accurate numeric information needed for statistical analysis. It is also more likely that interesting but unexpected results will be missed.

If little is known about who will profit most from using a system (or who will use it in a particular way), it is also important to collect data from a wide range of participants and to collect a wide range of demographic information.

With respect to measuring the effects, it would also be useful to develop a range of measures, unless there are clear theoretical or empirical a priori reasons for concentrating on a specific measure. If a particular type of learning is the goal of the system design, it would be useful to have some direct learning measure (such as recall, or problem-solving ability), and also to include a few indirect measures (transfer, metacognitive strategies via verbal protocols) or a delayed test condition. With any given learning paradigm it is likely, that differential effects will emerge for these different measures (e.g., Schmidt and Bjork, 1992). Spreading measurement techniques in this way may help to discover where exactly the beneficial effects of a system may lie.

The same tactic would also apply to other types of measurements, for example measuring the quality of collaboration. As long as the hypotheses are vague, it makes sense to collect a broad range of information to make sure that no interesting effects are overlooked.

To analyze these data, correlative data analysis techniques, such as standard regression techniques, correlations, and factor analyses can be used. Even though direct causal interpretations are not possible with these methods, they will help to identify interesting relationships in the data that can then be researched further with controlled designs. Unfortunately, correlative data analysis techniques can only be used with a relatively large number of subjects. Reeves (1993) suggests similar types of analyses techniques, and Shute (1993) provides a great example for the advantages of using correlations of actual use of functionality to explain performance.

In her study, Shute (1993) compared two different versions of an intelligent tutoring system that teaches electric circuit design. In both versions, students solve problems that are selected by the tutor based on previous problem-solving performance. Students also have access to a window with definitions of terms (e.g., 'voltage'). Another window gives access to simulations of circuits that can be used to take measurements or to experiment in the domain to extract additional knowledge. These two windows are accessible to students using both versions. The single difference between the two versions is the type of feedback that is given to students after a problem solution has been submitted to the system. In the rule-application version students are notified about mistakes, and the feedback also includes the rules that had to be applied but were violated. In the rule-induction version, students are only told whther or not they solved the problem correctly, and students are left to induce the rules themselves.

In her experiment, Shute collected data from over 300 subjects over seven days. She collected a range of data describing the subject population, such as demographic data, pretests, determining declarative and procedural, domain relevant background knowledge, and cognitive ability measures, such as working memory capacity and information processing speed. Furthermore she collected, for each subject, the frequency of use of the declarative and procedural exploratory features of the system. In order to measure the learning effects, she combined several posttests into one learning score, and she defined learning efficiency over the time needed to complete the tutor, and the number of problems needed to reach criterion (the more mistakes subjects made, the more problems were presented to them). The core of the statistical analyses are three multiple regressions, using performance on the combined posttest score, efficiency measured by time on tutor, and efficiency measured by -number of problems solved as the three dependent variables. The independent variables were posttest scores, version (induction vs. application), and the frequencies of using the declarative and procedural exploratory features of the system. Without going into the details of these analyses, Shute found a range of interesting results. Overall, subjects performed better in the rule application version, where they did not have to induce the rules of the domain. Secondly, the more the subjects made use of the declarative exploration feature, the better their performance overall. Procedural exploration, on the other hand, seemed to be related to poorer performance. However, there was also an interesting interaction in that procedural exploration in the rule-induction environment was related to improved performance, whereas it was related to depressed performance in the rule application version of the system.

These types of results, coming from a multiple regression analysis don't necessarily lead to clear causal interpretations, but they suggest interesting hypotheses to follow up in further studies or to consider in the design of systems: Students who like to explore more are not generally better learners, but they might be better suited to learn in a open environment with little specific guidance; procedural exploration may actually hinder performance when all necessary information can be found in the immediate performance; exploration of declarative and procedural aspects may have different effects on acquisition of requisite knowledge.

In a simple study using one dependent measure and one independent measure (version (a) vs. version (b)) this important information would have been lost, and no explanation for the differences in performance in the two studies could have been offered.

7. The Value of Theory or Hypothesis Guided Design

The most critical lesson from constructivist theories is that computer educational tools should treat users as active agents in the learning process, that deep learning requires situations where learners have to construct knowledge for themselves. You have to really own the new knowledge to be able to use it. Computers can make available a range of valuable tools to aid difficult conceptualizations and to encourage multifaceted learning (e.g., by transforming a textual description of a process into a dynamic display). Further, they can also provide supports for higher level, analytic thinking processes (e.g., by prompting the learner where to search for relevant ideas). However, both roles depend crucially on careful task analysis to select the right tool for a particular learner and a particular task.

Beyond their usefulness for individual learners, multimedia education systems may function as instruments of change in the classroom. Many of the programs we have reviewed are designed for or certainly are suitable for working in collaborative groups, ranging from two-to-three persons to a whole classroom and thus serve to break down the communication barriers between teacher and students that characterize so many didactic settings and that seriously interfere with getting students to think reflectively, to analyze ideas, to criticize and justify opinions.

We end this review with a plea for the use of theory in the design of instructional systems, including multimedia instructional systems. Both our general discussion and the analyses of particular systems and their performance suggest that it is impossible or at least very difficult and resource consuming to test and evaluate each component of a system in the field. Laboratory studies have their own limitations. Of course, we cannot do without either one. We need the field studies, the careful evaluations, however expensive and time consuming they might be. Further progress in our understanding of these systems depends on such work. But we also need quick and practical alternatives.

The theory of learning sketched in these pages can function as a guide for many design tasks. It will not be an infallible guide because, as it is obvious from the preceding discussions, learning is a very complex process, and its outcome depends on such a large number of factors that we can never hope to control them all. However, if one has a general understanding of the principles of learning and instruction as outlined here, as well as an appreciation of the nature of the learning task involved, one ought to do quite well in designing instructional systems according to those principles as a first approximation. Then, relying on the kind of informal observation and evaluation that is always possible, adjustments can be made, and hopefully, with a few iterations, a usable system can be developed. Nevertheless, a thorough task analysis is required in order to determine what type of learning is involved and when and what type of multimedia support would be most beneficial. As we have discussed above, tasks that are merely procedural versus those that can be learned in an apprenticeship situation, or by telling a story or showing a movie, each make their own kinds of demands.

What we have emphasized here are tasks requiring rational thought and symbolic representation. These tasks are the most difficult ones for students of all ages and educational levels; these are the ones that many instructional systems must focus upon. The principles of learning described above, and the systems that we described which embody these principles in one way or another, can be a guide to the design of such instructional systems.

repetition and reinforcement, or by active construction of meaning. All of these are perfectly good topics for the psychologist studying learning, but these different forms of learning must be differentiated for instructional purposes. Each requires a different set of learning principles.

The lower forms of learning that people share with other animals - habit formation through conditioning, motor learning through practice, incidental learning through experience, apprenticeship through expert guidance - really need not concern us very much, for people are very good at it. If the instructor sets up opportunities for learning, provides appropriate feedback, and most of all, makes sure that the learner remains sufficiently motivated to keep at the task, instruction will succeed. This does not mean that such learning is easy or simple, and that it does not play a great role in human life both inside and outside of school, nor that there are no principles of learning worth discussing that apply to such tasks. The behavior acquired can be very complex, very difficult, and a good trainer/instructor can make all the difference in the world. However, such learning is not what we are focusing on in the present report. Neither are we concerned with another type of human learning - learning from stories. Much of what people know is transmitted by stories, oral or written or visually presented. Indeed, human cultures could not exist without this way of teaching its children, for it is a primary vehicle for transmitting a culture's acquired knowledge and values. Students do learn well from stories and from films, but once again, knowing how to tell an effective story or how to make a successful film is by no means a trivial task.

Even though informal, everyday learning proceeds rather easily and naturally, and despite the fact that we know a lot today about how to improve students' learning, school learning remains difficult, effortful, and all too often is not successful. While schools teach a variety of things, the main reason for formal schooling in a culture has to do with acquiring higher forms of understanding which require abstract, symbolic thought. Symbolic thinking and reasoning processes are difficult, they do not come easily, and students need a great deal of support, guidance and encouragement in order to sustain their motivation and effort.

We begin the article with a general discussion of the nature of cognition in order to clarify our focus here on formal, symbolic thinking and learning. We then present a set of learning principles, which are justified and elaborated by means of an overview of the cognitive research literature from which they are derived. Afterwards we describe several multimedia educational systems to give some idea of how some or most of these principles have been implemented. We then describe how such systems have been evaluated from the standpoint of our list of learning principles and conclude the paper with a discussion of open questions for future research.

References

Adams, S. T., & diSessa, A. A. (1991). Learning by "cheating": Students' inventive ways of using a Boxer motion microworld. Journal of Mathematical Behavior, 10, 79-89.

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. Psychological Review, 94, 192-210.

Anderson, J. R. (1992). Intelligent tutoring and high school mathematics. Proceedings of the Second International Conference on Intelligent Tutoring Systems. Montreal.

Anderson, J. R., Conrad, F. G., & Corbett, A. T. (1989). Skill acquisition and the LISP tutor. Cognitive Science, 13, 467-506.

Bangert-Downs, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. R. (1991). The instructional effect of feedback in test-like events. Review of Educational Research, 61(2), 213-238.

Battig, W. F. (1979). The flexibility of human memory. In L. S. Cermak. &. F. I. M. Craik (Eds.), Levels of processing in human memory (pp. 23-44). Hillsdale, NJ: Erlbaum.

Beck, I. L., McKeown, M. G., Sinatra, G. M., & Loxterman, J. A. (1991). Revising social studies text from a text-processing perspective: Evidence of improved comprehensibility. Reading Research Quarterly, 27, 251-276.

Beck, I. L., McKeown, M. G., Worthy, J., Sandora, C., & Kucan, L. (in press). Questioning the author: A year-long implementation to engage students with text. Elementary School Journal.

Bereiter, C. (1994). Constructivism, socioculturalism, and Popper's World 3. Educational Researcher, 23(7), 21-23.

Bereiter, C., & Scardamalia, M. (1987). The psychology of written composition. Hillsdale, NJ: Erlbaum.

Bransford, J. D., Stein, B. S., Vye, N. J., Franks, J. J., Auble, P. M., Mezynski, K. J., & Perfetto, B. A. (1982). Differences in approaches to learning: An Overview. Journal of Experimental Psychology: General, 111(4), 390-398.

Britton, B. K., & Gulgoz, S. (1991). Using Kintsch's model to improve instructional text: Effects of inference calls on recall and cognitive structures. Journal of Educational Psychology, 83, 329-345.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In J. F. Flavell &. E. M. Markman (Eds.), Handbook of child psychology: Child development (pp. 515-629). New York, NY: Wiley.

Brown, A. L., & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), Classroom lessons: Integrating cognitive theory and classroom practice (pp. 229-270). Cambridge, MA: MIT Press.

Brown, A. L., & Day, J. D. (1983). Macrorules for summarizing text: The development of expertise. Journal of Verbal Learning and Verbal Behavior, 22, 1-14.

Brown, A. L., Day, J. D., & Jones, R. S. (1983). The development of plans for summarizing texts. Child Development, 54, 968-979.

Brown, A. L., & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), Knowing, learning, and instruction: Essays in honor of Robert Glaser (pp. 393-451). Hillsdale, NJ: Erlbaum.

Brown, A. L., & Reeve, R. A. (1987). Bandwidths of competence: The role of supportive context in learning and development. In L. S. Liben (Ed.), Development and learning: Conflict or congruence? (pp. 173-223). Hillsdale, NJ: Erlbaum.

Bruce, B. C., & Rubin, A. (1993). Electronic quills: A situated evaluation of using computers for writing in classrooms. Hillsdale, NJ: Erlbaum.

Bruner, J. S. (1986). Actual minds, possible worlds. Cambridge, MA: Harvard University Press.

Bulkeley, W. M. (1995, November 13). Back to School. Wall Street Journal [Special section of the state of technology in the nation's schools], R1-31.

Burton, R. R., Brown, J. S., & Fischer, G. (1984). Analysis of skiing as a success model of instruction: Manipulating the learning environment to enhance skill acquisition. In B. Rogoff & J. Lave (Eds.), Everyday cognition: Its development in social context (pp. 139-150). Cambridge, MA: Harvard University Press.

Chall, J. S. (1994). What students were reading 100 years ago. American Educator, 18(2), 26-33.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser R. (1989). How students study and use examples in learning to solve problems. Cognitive Science, 13, 145-182.

Chi, M. T. H., de Leeuw, N., Chiu, M-H, & LaVancher, C. (1994). Eliciting self-explanations improves understanding. Cognitive Science, 18, 439-477.

Clark, Richard E. (1983). Reconsidering research on learning from media. Review of Educational Research, 53, 445-459.

Cognitive Technology Group at Vanderbilt (1992). The Jasper experiment: An exploration of issues in learning and instructional design. Educational Technology Research and Development, 40(1), 65-80.

Cognitive Technology Group at Vanderbilt (1993). Anchored instruction and situated cognition revisited. Educational Technology, (March 1993), 52-69.

Cognitive Technology Group at Vanderbilt, (1994). From visual word problems to learning communities: Changing conceptions of cognitive research. In K. McGilly (Ed.), Classroom lessons: Integrating cognitive theory and classroom practice (pp. 157-200). Cambridge, MA: MIT Press.

Coleman, E. B. (1994). Effects of explanation vs. summarization of science texts on memory and inference. Paper presented at the meetings of the Winter Text Conference, Jackson, WY.

Collins, A., Brown, J. S., & Holum A. (1991). Cognitive apprenticeship: Making thinking visible, American Educator, 15(3), 6-11.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading writing, and mathematics. In L. B. Resnick (Ed.), Knowing, learning, & instruction (pp. 453-494). Hillsdale, NJ: Erlbaum.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychologies. American Psychologist, 30, 1-13.

Dansereau, D. F. (1978). The development of learning strategies curriculum. In J. H. F. O'Neil (Ed.), Learning strategies. New York, NY: Academic Press.

Dansereau, D. F. (1985). Learning strategy research. In J. W. Segal, S. Chipman, & R. Glaser (Eds.), Thinking and learning skills. Vol. 1: Relating instruction to research (pp. 209-239). Hillsdale, NJ: Erlbaum.

Dewey, J. (1938). Progressive education and the science of education. Washington, DC: Progressive Education Association.

diSessa, A. A., & Minstrell, J. (in press). Cultivating conceptual change with Benchmark Lessons. In J. G. Greeno (Ed.), Thinking practices.

Donald, M. (1991). Origins of the modern mind. Cambridge, MA: Harvard University Press.

Druckman, D., & Bjork, R. A. (Eds.). (1994). Learning, remembering, believing: Enhancing human performance. Washington, D. D.: National Academy Press.

Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis. Cambridge, MA: MIT Press.

Ferstl, E., & Kintsch,W. (1991). Learning from text: Structural knowledge assessment in the study of discourse comprehension. Unpublished manuscript.

Fischer, G., McCall, R., & Morch, A. (1989). JANUS: Integrating hypertext with a knowledge-based design. Hypertext '89 Proceedings, (pp. 105-117). Pittsburgh, PA: Association of Computing Machinery.

Fitts, P.M. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), Categories of Human Learning. New York: Academic Press, 1964.

Foltz, P. W. (1992). Reader's comprehension and strategies in linear text and hypertext. Unpublished Ph.D. dissertation. University of Colorado.

Gallistel, C. R. (1990). The organization of learning. Cambridge, MA: MIT Press.

Garner, R. (1988). Metacognition and reading comprehension. Norwood, NJ: Ablex Publishing Company.

Gick, M. L., and Holoyak, K. J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.

Glaser, R. &. Bassok, M. (1989). Learning theory and the study of instruction. Annual Review of Psychology, 40, 631-666.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. Journal of Experimental Psychology: Learning, Memory and Cognition, 11, 501-518.

Gräsel, C., Mandl, H., Fischer, M, Gärtner, R. (1994). Vergebliche Designermüh? Interaktionsangebote in problemorientierten Computerlernprogrammen. (Tech. Rep. No. 38) Ludwig-Maximilans-Universität München, Institut für Pädagogische Psychologie und Empirische Pädagogik, Munich, Germany.

Greeno, J. G. (1994). Gibson's affordances. Psychological Review, 101, 336-342.

Hansen, J., & Pearson, P. D. (1983). An instructional study: Improving the inferential comprehension of good and poor fourth-grade readers. Journal of Educational Psychology, 75, 821-829.

Hoban, C. F. (1958). Research on media. AV Communications Review, 6, 169-178.

Hovland, C., Lumsdaine, A. A., & Sheffield (1949). Experiments on mass communication. Princeton, N.J.: Princeton University Press.

Interactive Educational Systems Design (1995). Report on the effectiveness of technology in schools 1990-1994. (Tech. Rep.) Software Publishers Association.

Keefe, D. E., & McDaniel, M. A. (in press). The time course and durability of predictive inferences. Journal of Memory and Language.

Kintsch, E. (1990). Macroprocesses and microprocesses in the development of summarization skill. Cognition and Instruction, 7(3), 161-195.

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

Kintsch, W., & Welsch, D. (1991). The construction-integration model: A framework for studying memory for text. In W. E. Hockley & S. Lewnadowsky (Eds.), Relating theory and data: Essays on human memory in honor of Bennet B. Murdock (pp. 367-385). Hilldale, NJ: Erlbaum.

Koedinger, K. R., and Anderson, J. R (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. Cognitive Science, 14, 511-550.

Landauer, T. K., Egan D. E. , Rewmde, J. R., Lesk, M. E., Lochbaum, C. C. & Ketchum, R. D. (1993). Enhancing the usability of texts through computer delivery and formative design evaluation. In C. M. A. Dillon & J. Richardson (Eds.), Hypertext: A psychological perspective (pp. 71-136). New York NY: Ellis Horwood.

Larkin, J. H., Heller, J. I., & Greeno, J. G. (1980). Instructional implications of research on problem solving. In W. J. McKeachie (Ed.), Cognition, college teaching, and student learning (pp. 51-66). San Francisco, CA: Jossey-Bass.

Larkin, J. H. (1989). What kind of knowledge transfers? In L. B. Resnick (Ed.), Knowling, learning, and instruction: Essays in honor of Robert Glaser (pp. 283-305). Hillsdale, NJ: Erlbaum.

Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows & H. A. Houghton (Eds.), The psychology of illustration (pp. 51-85). New York: Springer.

Levin, J. R., & Mayer, R. E. (1993). Understanding illustrations in text. In A. Woodward, B. K. Britton & M. Binkley (Eds.), Learning from textbooks. (pp. 95-114). Hillsdale, NJ: Erlbaum.

Mannes, S., & Kintsch, W. (1987). Knowledge organization and text organization. Cognition and Instruction, 4, 91-115.

Mayer, R. E. (1980). Elaboration techniques that increase the meaningfulness of technical text: An experimental test of the learning strategy hypothesis. Journal of Educatinal Psychology, 72(6), 770-784.

Mayer, R. E. (1989). Models for understanding. Review of Educational Research, 59(1), 43-64.

Mayer, R. E. (1994). Visual aids to knowledge construction: Building mental representations from pictures and words. In W. Schnotz & R. W. Kulhavy (Eds.), Comprehension of graphics. Amsterdam: Elsevier.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? Journal of Educational Psychology, 82(4), 715-726.

McNamara, D. S., & Kintsch, W. (in preparation). Learning from a history text: Effects of prior knowledge and text coherence.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (in press). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. Cognition and Instruction.

Minstrell, J. (1989). Toward science for understanding. In L. B. Resnick &. L. E. Klopfer (Eds.), Toward the thinking curriculum: Current cognitive research (pp. pp. 129-149). Alexandria, VA: Association for Supervision and Curriculum Development.

Mishkin, M., & Petri, H. L. (1984). Memories and habits: Some implications for the analysis of learning and retention. In L. R. Squire &. N. Butters (Eds.), Neuropsychology of memory (pp. 287-297). New York, NY: Guilford Press.

Nathan, M. J., Bransford, J., Brophy, S., Garrison, S., Goldman, S R., Kantor, R. J., Vye, N., & Williams, S. (1994). Multimedia journal articles: Promises, pitfalls, and recommendations. Educational Media International, 31 (4), 265-273.

Neisser, U. (in press). Multiple systems: A new approach to cognitive theory. The European Journal of Cognitive Psychology.

Nelson, K. (in press). Language in cognitive development: The emergence of the mediated mind. New York, NY: Cambridge Unviersity Press.

Norman, D. (Ed.) (1993). Situated action [Special issue] Cognitive Science, 1.

O'Hara, K. P., & Payne, S. J. (submitted). Cost of operations affects planfulness of problem solving.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. Cognition and Instruction, 1(2), 117-175.

Papert, S. (1993). The children's machine: Rethinking school in the age of the computer. New York, NY: Basic Books, Harper Collins.

Pennington, N, Nicolich, R., & Rahm, J., (1995). Transfer of training between cognitive subskills: Is knowledge use specific? Cognitive Psychology, 28, 175-224.

Pennington, N., & Nicolich, R. (1991). The transfer of training between progamming subtasks: Is knowledge really use specific? In J. Koenemann-Belliveau, T. G., Moher, & S. P. Robertson (Eds.), Empirical studies of programmers (pp. 156-176). Norwood, N.J: Ablex.

Rauenbusch, F., & Bereiter, C. (1991). Making reading more difficult: A degraded text microworld for teaching reading comprehension strategies. Cognition and Instruction, 8(2), 181-206.

Reeves, T. C. (1986). Research and evaluation models for the study of interactive video. Journal of Computer-Based Instruction, 13, 102-106.

Reeves, T. C. (1990). Redirecting evaluation of interactive video: The case for complexity. Studies in Educational Evaluation, 16, pp. 115-131.

Reeves, T. C. (1993). Research support for interactive multimedia: Existing foundations and new directions. In J. W. C. Latchem & L. Henderson-Lancett (Eds.), Interactive multimedia: Practice and promise (pp. 79-96). London: Kogan Page.

Resnick, L. B. (1989). Introduction. In L. B. Resnick (Ed.), Knowing, learning, and instruction: Essays in honor of Robert Glaser (pp. 1-24). Hillsdale, NJ: Erlbaum.

RobertJan Simons, P. (1982). Concrete analogies as aids in learning from text. In A. Flammer & W. Kintsch (Eds.), Discourse processing (pp. 462-471). Amsterdam: North-Holland Publishing Co.

Ross, B. H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 456-468.

Rousseau, J. J. (1979). Emile . Basic Books.

Samuels, S. J. (1970). Effects of pictures on learning to read, comprehension and attitudes. Review of Educational Research, 40, 397-407.

Scardamalia, M., & Bereiter, C. (1993). Technologies for knowledge-building discourse. Communications of the ACM [Special issue: Technology in education] 367(5), 37-40.

Scardamalia, M., Bereiter, C., & Lamon, M. (1994). The CSILE project: Trying to bring the classroom into World 3. In K. McGilley (Ed.), Classroom lessons: Integrating cognitive theory and classroom practice (pp. 201-228). Cambridge, MA: MIT Press.

Scardamalia, M., Bereiter, C., McLean, R. S., Swallow, J., & Woodruff, E. (1989). Computer-supported intentional learning environments. Journal of Educational Computing Research, 5(1), 51-68.

Scardamalia, M., Bereiter, C., & Steinbach, R. (1984). Teachability of reflective processes in written composition. Cognitive Science, 8, 173-190.

Schramm, W. (1977). Big media little media. Beverly Hills, CA: Sage.

Schmidt, R. A. (1988). Motor control and learning: A behavioural emphasis. Champaign, IL: Human Kinetics Publishers.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. Psychological Science, 3(4), 207-217.

Schooler, L. J., & Anderson, J. R. (1990). The disruptive potential of immediate feedback. Proceedings of the Cognitive Science Society. Hillsdale, NJ: Erlbaum.

Shapiro, A. M. (submitted). Learning from hypermedia: The effect of information structure on complex concept acquisition.

Shute, V. J. (1993). A comparison of learning environments: All that glitters... In S. P. Lajoie & S. J. Derry (Eds.), Computers as cognitive tools (pp. 47-73). Hillsdale, NJ: Erlbaum.

Singley, M. K., and Anderson, J. R. (1989). The transfer of cognitive skill. Cambridge, MA: Harvard University Press.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1991). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill structured domains. Educational Technology, 31, 224-33.

Spiro, R. J., & Jehng, J. C. (1990). Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix and R. J. Spiro (Eds.), Cognition, education, and multimedia: Exploring ideas in high technology (pp. 163-205). Hillsdale, NJ: Erlbaum.

Spiro, R. J., Vispoel, W. P., Schmitz, J. G., Samarapungavan, A., & Boerger, A. E. (1987). Cognitive flexibility and transfer in complex content domains. In B. K. Britton &. S. M. Glynn (Eds.), Executive control processes in reading (pp. 177-199). Hillsdale, NJ: Erlbaum.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis of findings with rats, monkeys, and humans. Psychological Review, 99, 195-231.

Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. Annual Review of Psychology, 44, 453-495.

Svendsen, G. B. (1991). The influence of interface style on problem solving. International Journal of Man-Machine Studies, 35, 379-397.

Time Magazine (1995 Spring). Welcome to Cyberspace [Special issue], 145(12), 49-51.

Van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. New York: NY: Academic Press.

Van Lehn, K. (1988). Toward a theory of impasse driven learning. In H. Mandl & A. Lesgold (Eds.), Learning issues for intelligent tutoring systems (pp. 19-41). New York, NY: Springer.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. (J. Cole, V. John-Steiner, S. Scribner, & E. Souberman Eds. & Trans.). Cambridge, MA: Harvard University Press.

Weinert, F. E., & Helmke, A. (1995). Learning from wise mother nature or big brother instructor: The wrong choice as seen from an educational perspective. Educational Psychologist, 30(3), 135-142.

Weinstein, C. E., Götz, E. T., Alexander, P. A. (Eds.) (1988). Learning and study strategies.

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), Handbook of research on teaching (pp. 315-327). New York, NY: Macmillan.

Woodward, A. (1993). Do illustrations serve an instructional purpose in US textbooks? In A. Woodward, B. K. Britton & M. Binkley (Eds.), Learning from textbooks (pp. 115-134). Hillsdale, NJ: Erlbaum.

2. The Nature of Cognition

Anyone who tries to gain an overview of the contemporary literature on learning and instruction will find that it is an amazingly contentious and fractionated area. One would think that after 100 years of the psychological study of learning there would be some agreement about the basics, at least. On the contrary, there appears to exist a Babylonian confusion today at the level of basic approaches to learning and instruction. Once we get beyond that point and look into each particular approach by itself, things appear much more orderly. But at the level of basic outlook, confusion reigns.

Therefore, we begin this review of the literature with a general discussion about the nature of cognition, which may seem like a detour but is necessary to avoid confusion. There are many successful learning systems in use which operate on quite different principles than the ones discussed here. Our focus is on formal school learning. However, people learn - and think - in many different ways, to which the principles listed above do not apply. The brief discussion below on forms of cognition should help to clarify what these ways are, how they interact, and why we chose to focus on one particular type of learning and instruction.

One source of confusion regarding learning and instruction is the fact that earlier research on conditioning and habit formation in the behavioristic tradition has never been reconciled with the cognitive approach, which looks at learning as a process of meaning and knowledge construction. The cognitive approach has simply shifted the discourse to a new level and taken over the spotlight, but habit learning and skill instruction remains alive and well in many settings. More recently the cognitive/constructivist approach has itself been challenged by demonstrations of situated cognition and learning through direct action, revealing a number of significant phenomena that cannot be accounted for by symbolic cognitive theories. In some sense, situativity theory takes us all the way back to non-representational behaviorism, though the former's emphasis on social and cultural factors in learning adds a completely new and significant dimension. Communication between these approaches is generally poor, as attested by the special issue of Cognitive Science dedicated to the symbol-vs.-situation controversy (Norman, 1993).

One cannot state principles of learning without resolving these confusions. Therefore we briefly discuss a framework that distinguishes different types of learning and argue that formal pedagogy must emphasize one particular form of learning and knowing - the construction of symbolic mental models. The set of learning principles presented in the following section focuses on that particular goal. In articulating this framework we hope to avoid fruitless discussions and counter examples to our proposed principles that involve different types of learning and cognition.

Technical Report 96-01

University of Colorado at Boulder

Institute of Cognitive Science

Muenzinger Psychology Building
Campus Box 344
Boulder, Colorado 80309-0344
(303) 492-5063
FAX: (303) 492-7177
email: cogsci@psych.colorado.edu
URL: http://psychwww.colorado.edu/ics/