J. EDUCATIONAL COMPUTING RESEARCH, Vol. 18(2) 123-145, 1998

DESIGNING A HYPERMEDIA TOOL FOR LEARNING ABOUT CHILDREN'S MATHEMATICAL COGNITION

MATTHEW J. KOEHLER
RICHARD LEHRER
University of Wisconsin–Madison

ABSTRACT

We designed a hypermedia tool for helping preservice teachers learn about the growth and development of children's mathematical reasoning. Design of the tool was guided by a small set of design principles derived from consideration of research in cognitive science. These included "criss-crossing" the conceptual landscape and developing navigational tools that made this landscape visible. We compared learning in text and hypermedia environments by employing a novel single-subject methodology that afforded an economical means of assessment. Results indicated that participants learned significantly faster using the hypermedia system than they did using text-based materials. Secondary analyses suggest that this finding could be attributed to several elements of the design, including the learners' access to examples, an interface that was flexible enough to meet readers' changing experience and goals, and the conceptual criss-crossing embedded in the structure of the system.

INTRODUCTION

Hypermedia expands the "writing space" of authors because nodes of information can be configured in any way and traversed in any order [1]. However, it is not clear whether or not this expansion of the writing space is always accompanied by a corresponding expansion of opportunities to learn [2]. Accordingly, we designed, developed, and tested a hypermedia system that was intended to assist preservice teachers learn about the growth and development of young children's mathematical reasoning. We view our cycle of design and assessment as a testbed for the prospects and pitfalls of hypermedia systems for learning. In the first section of this article, we describe a small number of design principles for learning with hypermedia, based on research in cognitive science. In the second

doi: 10.2190/HFLM-ER1P-6J25-V0XD
http://baywood.com

section, we describe how these general design principles were realized in the design of the system to help preservice teachers learn about children's mathematical reasoning. We conclude with a report of an experiment employing single-subject methodology to compare preservice teachers' learning in text and hypermedia environments.

DESIGN PRINCIPLES

Unlike traditional texts, hypermedia and hypertext documents often have an underlying non-linear ("web-like") structure. The promise of hypermedia, from the vantage point of its potential to scaffold learning, consists in large measure of its capacity to represent the complexity of the domain. Moreover, hypermedia documents afford learner interaction with the domain. Yet, learning with hypermedia systems seems to be sensitive to a number of factors, including the structure of the domain and the design of the interface [3, 4]. Hypermedia seems well-suited to the learning of complex, ill-structured domains, but interfaces are often designed in ways that place heavy navigational burdens on readers. Consequently, much of the potential advantage of hypermedia is often dissipated.

We turned to the emerging research in cognitive science for insights about how to best exploit the potential of hypermedia for designing effective environments for learning. Although cognitive science is multidisciplinary and still evolving, nonetheless cognitive research in areas like problem solving, tutoring, and the design of intelligent systems has progressed sufficiently to suggest heuristics for the design of cognitive tools [5]. We express this relationship between learning with hypermedia and insights garnered from cognitive science as a set of design principles.

Use Non-Linearity to "Criss-Cross" the Landscape

Cognitive Flexibility Theory [6] maintains that for advanced knowledge acquisition in complex, ill-structured domains, learners must "criss-cross" the conceptual landscape of that domain. Sites in a landscape (cases, or concepts in the knowledge domain) must be revisited from different directions, and thought about from different perspectives or "lenses." The additional cognitive load [7] introduced by the non-linear, multi-dimensional structure of hypermedia may serve no purpose in simple, well-defined domains. However, hypermedia's node-link structure is well-suited to provide "criss-crossing" of ill-structured domains.

Make the Structure of the Domain Visible

When expert and novice representations of domain knowledge are characterized, experts demonstrate more principled and more numerous connections between individual pieces of knowledge than their novice counterparts [8]. To help novices move toward a better representation of a domain, the linking structure of hypermedia documents should approximate the connections and relationships in the domain. In other words, the representational landscape should explicitly model the conceptual landscape of the domain (to the extent that this landscape is accessible and conventional). To make these connections and associations explicit, hypermedia documents should clearly signal the presence and nature of links. Hence, links should be *typed* so that readers can infer the reason for connections, and they should be clearly signaled (e.g., iconic buttons or typographical cued text) so that readers have easy access to these connections. Hypermedia systems should also include navigational tools that explicitly show associational structure to readers (e.g., graphical browsers).

Make Navigation Easy

Although the non-linearity of hypermedia allows new "writing space" for authors, non-linearity also puts additional demands upon its' readers [9]. Unlike readers of traditional texts, where authors sequence information, readers of hypermedia are required to select the learning sequence. This sequencing process, called "navigation," is demanding of working memory [10], a concern because working memory traditionally plays an important role in comprehension of text [11]. Hence, when the resource demands of navigation and learning jointly exceed working memory capacity, readers can become confused and "lost in hyperspace." To mitigate this problem, hypermedia systems should make navigation as easy as possible. As mentioned previously, whenever possible, navigational tools should make the structure of the domain more visible.

Readers have a lot of experience with the conventions and tools used in traditional, linear texts. Hypermedia systems should take advantage of this familiarity and extend these conventions to an "electronic book" metaphor to its readers. For example, readers traditionally expect to find a *table of contents* to start a document. This expectation should be fulfilled in hypermedia systems as well. Likewise, readers have familiarity and experience with *outlines, indexes, bookmarks*, and *advance organizers*—all of which may be used as navigational tools. When implemented electronically, these and related tools help readers chart a course in hyperspace by helping them find landmarks, remember their trails and goal-subgoal relationships, backtrack or retrace their steps, chose new information, and maintain connections between preceding and subsequent information. Moreover, navigational tools should be rendered in ways that display webs of association and conceptual neighborhoods, for it is just these forms of structure that can be made visible in hypermedia, but not in conventional text.

Navigation patterns are sensitive to a wide array of individual differences, including—domain expertise [12, 13], experience with the interface [14, 15], learning goal [16], learning style [17], spatial ability [4, 18], and locus of control [19]. For this reason, readers need access to a wide variety of navigational tools, so that the interface is aligned with individual differences and responsive to learners' changing needs as they gain experience with the system [4, 20].

Provide Learners Opportunities to Learn by Example

Goodman suggests that exemplification is a widely used mode of symbolization: the example possesses the properties that it refers to [21]. Examples are particularly important when properties to be learned cannot be explicitly stated, or when such properties are highly related and "criss-crossed" in the manner described previously. Moreover, examples often play a central role in case-based learning [22] and in helping learners construct relationships [23]. Hypermedia systems provide opportunities to employ video and audio examples, as well as more traditional examples based on text and illustration.

Layer Annotations and Examples

The act of navigation in hyperspace places additional demands upon the working memory of learners [9, 10]. With more memory tied up managing navigational tasks, less working memory can be devoted to other processes of learning. Hence, we suggest "layering" information to make it easier for readers to maintain relationships among associated concepts. For example, Black, Wright, Black, and Norman found that when definitions and the main text were presented on different screens, readers often had trouble relocating themselves in the main text [24]. In contrast, when definitions were layered (highlighted and presented on the same screen), readers accessed definitions more often and with less difficulty. Text and video examples should be layered as well, because examples only make sense in relation to the principles conveyed in the main text. That is, to claim that something is an "example" is to claim that it is an "example of something." For readers to make the connection between an example and a larger principle requires access to both.

Lessons Learned

Of course, principles of interface design established by instructional designers should not be forgotten when designing hypermedia tools for learning (see [25, 26]). For example, hypermedia systems should offer consistent visual cues to signal functionality (e.g., the purpose of buttons) and to provide feedback for reader actions (e.g., highlight buttons when they are pushed and signal screen transitions). Readers also need easy access to features and shortcuts, so that the most useful and frequently used features are visible and accessible.

DESIGN AND DEVELOPMENT OF HyperCGI

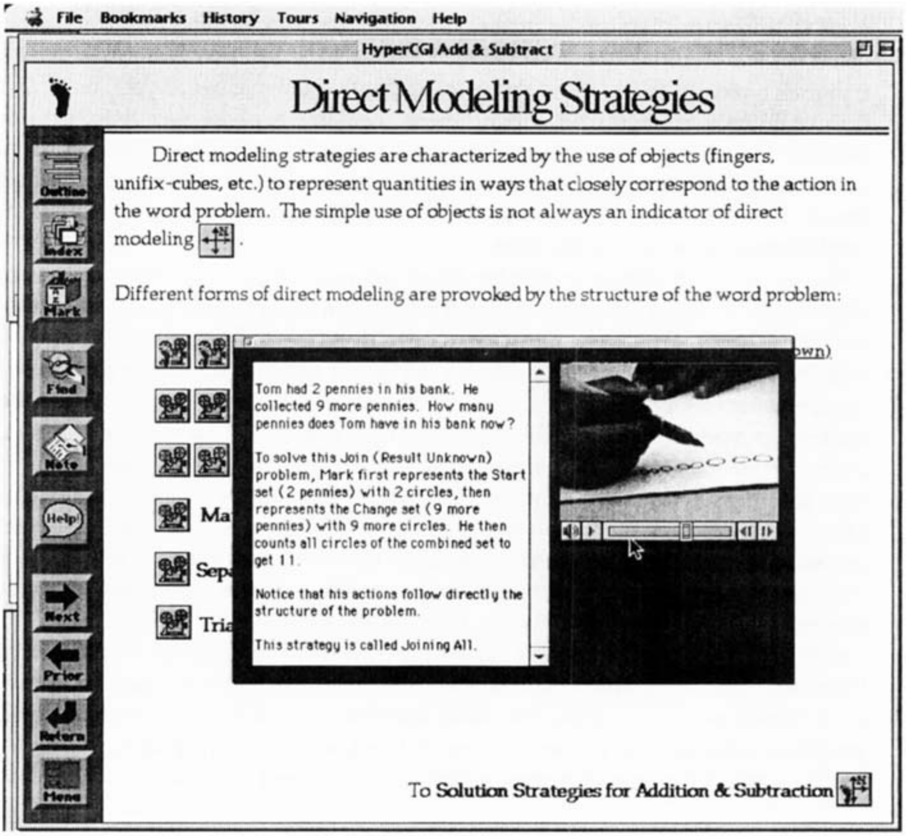
We used these design principles to guide the development of a hypermedia tool to help preservice teachers learn about the growth and development of children's mathematical reasoning. The system was designed to inform readers about Cognitively Guided Instruction (CGI), a program of professional development that aims to reform primary-grade mathematics education by helping teachers understand student thinking about arithmetic [27]. In CGI, teachers learn how young children typically think about the semantics of arithmetic word problems, and how children's strategies for solving these problems evolves over time. Consequently, the design of the hypermedia system focused on these two components of the program (the semantics of word problems and the strategies that children invent to solve them). Other aspects of CGI, like typical issues in classroom implementation, were not addressed.

We applied the design principles discussed previously to develop a system, HyperCGI, to foster the professional development of elementary school mathematics teachers. Presenting the CGI model in hypermedia form represented a significant design challenge because the domain is semantically rich, and many of the elements of the domain, such as the semantics of the word problems and the strategies children use to solve them, are best understood in relation. That is, many of the concepts are constituted as conceptual landscapes, not as single elements in isolation. For example, the nature of a child's solution strategy is often a consequence of the semantics of a word problem, so that teachers select certain classes of word problems to provoke the development of certain kinds of solution strategies. When completed, the resulting system contained sixty-six screens of information (see Figure 1 for an example screen). In addition, 165 annotations were provided in the form of text, video, and graphics. Ninety-six of these annotations provided readers access to thirty-two video-digitized episodes of children's problem solving. The remaining fifty-nine annotations were text and graphical elaborations. In the sections that follow, we describe how we implemented each of the cognitive design principles described previously.

"Criss-Crossing" the Landscape

The landscape of HyperCGI includes nodes devoted to the semantics of word problems (problem types), prototypical solution strategies invented by children (solution strategies), typical developmental trajectories, video episodes of children's problem solving, text examples, expert commentaries on children's actions, and information about diagnosis and assessment of students' mathematical learning. The conceptual landscape is criss-crossed by links associating different nodes in ways that outline their connections. Moreover, readers can browse the document with any of four "lenses" or viewpoints: problem-types, solution strategies, examples, and tours.

From the viewpoint of problem-type, the semantics of word problems provide the primary focus, and these are associated with relevant information about likely solution strategies and related examples. In another view, the primary perspective is solution strategies, with connections to different types of problems, developmental sequences, and examples. A third view emphasizes opportunities to learn by example. Under this view, information is always accompanied by either text or



Direct Modeling Strategies

Direct modeling strategies are characterized by the use of objects (fingers, unifix-cubes, etc.) to represent quantities in ways that closely correspond to the action in the word problem. The simple use of objects is not always an indicator of direct modeling.

Different forms of direct modeling are provoked by the structure of the word problem:

Tom had 2 pennies in his bank. He collected 9 more pennies. How many pennies does Tom have in his bank now?

To solve this Join (Result Unknown) problem, Mark first represents the Start set (2 pennies) with 2 circles, then represents the Change set (9 more pennies) with 9 more circles. He then counts all circles of the combined set to get 11.

Notice that his actions follow directly the structure of the problem.

This strategy is called Joining All.

To Solution Strategies for Addition & Subtraction

Figure 1. Example screen with a layered video example.

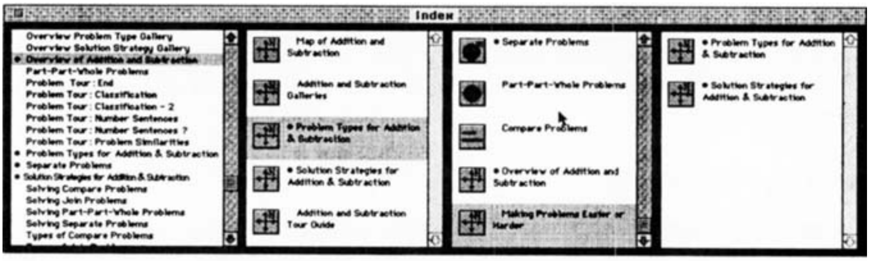
video examples; readers interact with *Galleries* that index a library of problem type and solution strategy examples. *Guided Tours* provide a fourth way of encountering the conceptual landscape of CGI. Tours allow readers to take a previously ordered path through the Web-like structure of the hypermedia document. The sequence of information was developed by a CGI expert (Tom Carpenter) to provide readers "ideal learning paths."

Making the Structure of the Domain Visible

HyperCGI uses typed iconic links to clearly signal connections in the domain. Link types are based upon the semantics of the domain (e.g., a class of solution strategy), and also are used to signal transitions in media (i.e., a camera icon to signal a video example). Graphical browsers are not used in HyperCGI, due to the problems of displaying such representations in large-scale hypermedia [9]. Instead, a dynamic *Index* tool shows a more localized representation of this structure (see Figure 2a). This Index displays an alphabetical listing of all the screens in the system. Clicking on any Index entry calls up the corresponding screen. The Index can also be expanded to show connections between screens. That is, clicking once on any Index entry creates a new list that contains each connected (linked) card, and icon representing the type of that connecting link. Entries in this secondary list can also be selected to show further connections. Structure is also made visible with a dynamic *Outline* tool (see Figure 2b). The levels of the Outline can be expanded or contracted by readers, so that hierarchical structure can be viewed in varying amounts of detail.

Navigation

To accommodate the diversity of tasks that readers must perform, as well as learner differences, HyperCGI provides a number of facilities to make navigation



a. Index

b. Outline

Figure 2. The *Index* and *Outline* tools.

as easy as possible for readers. The *Outline* and *Index* serve as navigational tools; readers can double-click on any entry to go to the target screen in the system. *Next* and *prior* buttons always allow the reader to get some next and previous screen in a predefined default linear order. Availability of next and prior seems especially important for novices [10, 28].

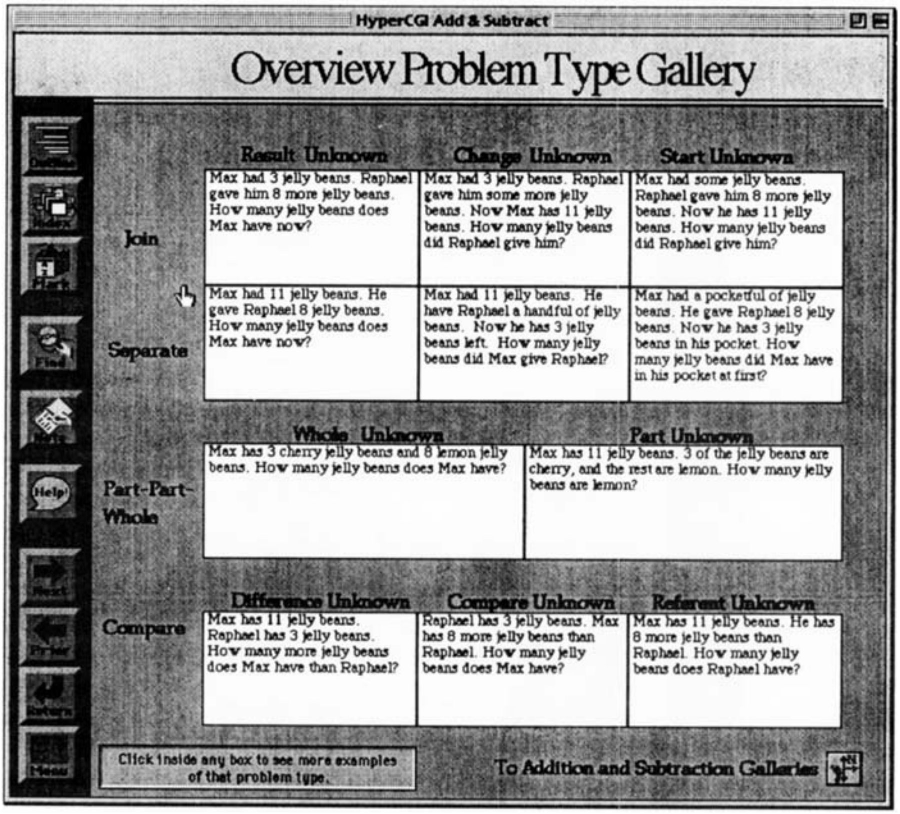
HyperCGI also provides several features to help readers manage their trails or histories in hyperspace. *Footprinting* uses visual cues to distinguish between new and old information. Screens that have been previously visited include a blue footprint icon in the upper left hand corner, while newly visited screens do not include this footprint (Figure 1 contains an example with such a footprint). Footprinting of screens is echoed in tools, so that, for example, typed links show an overlaid blue footprint when the destination has been visited, and both outline and index entries are marked if they have been visited. Furthermore, HyperCGI maintains a history of readers' trails, and provides readers opportunities to backup, or retrace their steps as needed. Consequently, the system always provides a *history* menu that lists the names of the most recently visited screens, allowing readers to retrace their navigational paths. Similarly, the *backtrack* button allows readers to retrace their paths one step at a time. Finally, *bookmarks* can also be created, so that readers later can recall landmarks in their exploration of hyperspace.

Learning by Example

HyperCGI is replete with text examples of problems and video examples of children's solution strategies. There are ninety-six annotation links containing access to the thirty-two video episodes of children's problem solving. These examples are embedded within individual screens to exemplify the principle in the text (see Figure 1 for an example). Because video often leaves it up to readers to discover what the video is a case or example of, a text explanation usually accompanies the video segment. This explanation summarizes the problem to be solved, the process the child used to solve the problem (in CGI language) and points out salient features of the segment. Figure 3 displays a gallery of text examples, arranged by problem types. In this gallery, readers can see examples of each of the problem types in CGI. Readers can click on any of the boxes to get a new example of that problem type. There is a similar gallery for solution strategies, organized by problem type and children's developmental level.

Layering

Layering allows connected information to be displayed on the same screen as the anchoring information. In HyperCGI, all annotations (e.g., expert commentary) and video examples always are displayed in layered pop-up windows (see Figure 1 for an example). Also, when deployed, the navigational tools (e.g., Index, Outline, Tours) are always displayed in a new layered window. When the



Overview Problem Type Gallery

	Result Unknown	Change Unknown	Start Unknown
Join	Max had 3 jelly beans. Raphael gave him 8 more jelly beans. How many jelly beans does Max have now?	Max had 3 jelly beans. Raphael gave him some more jelly beans. Now Max has 11 jelly beans. How many jelly beans did Raphael give him?	Max had some jelly beans. Raphael gave him 8 more jelly beans. Now he has 11 jelly beans. How many jelly beans did Raphael give him?
Separate	Max had 11 jelly beans. He gave Raphael 8 jelly beans. How many jelly beans does Max have now?	Max had 11 jelly beans. He have Raphael a handful of jelly beans. Now he has 3 jelly beans left. How many jelly beans did Max give Raphael?	Max had a pocketful of jelly beans. He gave Raphael 8 jelly beans. Now he has 3 jelly beans in his pocket. How many jelly beans did Max have in his pocket at first?

	Whole Unknown	Part Unknown
Part-Part-Whole	Max has 3 cherry jelly beans and 8 lemon jelly beans. How many jelly beans does Max have?	Max has 11 jelly beans. 3 of the jelly beans are cherry, and the rest are lemon. How many jelly beans are lemon?

	Difference Unknown	Compare Unknown	Referent Unknown
Compare	Max has 11 jelly beans. Raphael has 3 jelly beans. How many more jelly beans does Max have than Raphael?	Raphael has 3 jelly beans. Max has 8 more jelly beans than Raphael. How many jelly beans does Max have?	Max has 11 jelly beans. He has 8 more jelly beans than Raphael. How many jelly beans does Raphael have?

Click inside any box to see more examples of that problem type.

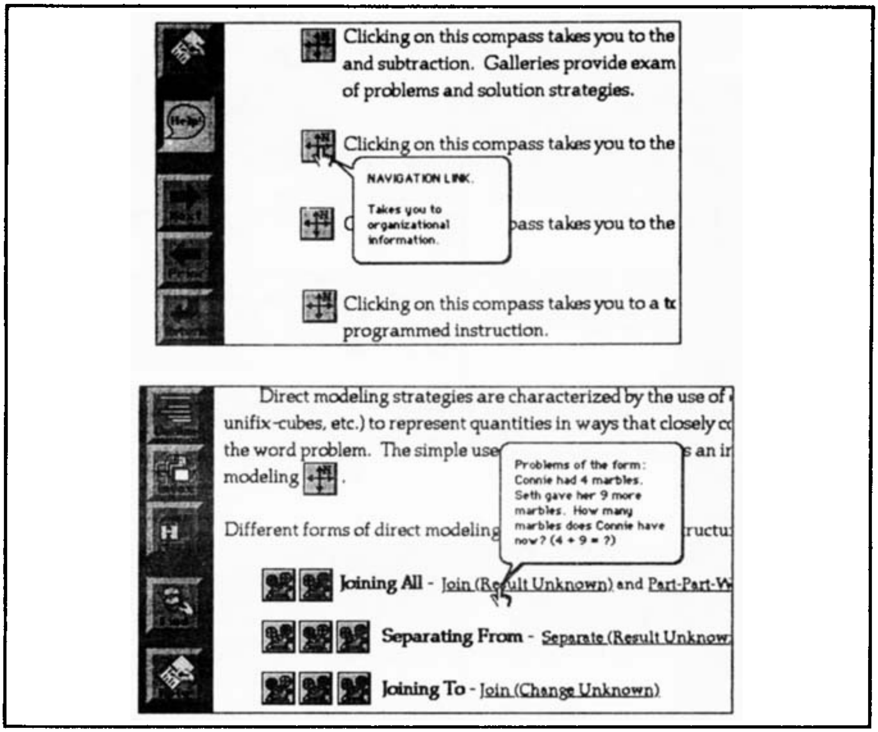
Figure 3. Overview problem type gallery.

reader clicks on an underlined word, a *short definition* is displayed as a memory aid (see Figure 4a). Context-sensitive *balloon help* provides readers with short help text for any object in the system (see Figure 4b).

During the design process, HyperCGI underwent several cycles of reader testing, so our implementation of principles was guided by the pragmatic concerns of constraints and affordances. For example, the location and grouping of buttons in the "toolbar" (see Figure 1) changed frequently in response to reader testing. After the final product was created, an experiment was conducted to test the effectiveness of the design, as we describe in the next section.

EXPERIMENTAL STUDY

We designed an experiment to test the effectiveness of HyperCGI as a tool for learning. Although many questions could be asked about different elements of the



Clicking on this compass takes you to the
and subtraction. Galleries provide exam
of problems and solution strategies.

Clicking on this compass takes you to the

NAVIGATION LINK.

Takes you to
organizational
information.

C
pass takes you to the

Clicking on this compass takes you to a t
programmed instruction.

Direct modeling strategies are characterized by the use of
unifix-cubes, etc.) to represent quantities in ways that closely c
the word problem. The simple use
s an i
modeling .

Problems of the form:
Connie had 4 marbles.
Seth gave her 9 more
marbles. How many
marbles does Connie have
now? (4 + 9 = ?)

Different forms of direct modeling
ructu

Joining All - Join (Result Unknown) and Part-Part-W

Separating From - Separate (Result Unknow

Joining To - Join (Change Unknown)

Figure 4. Layering features. (a) Context-sensitive Balloon Help. Here, help is turned on, and the reader has moved the mouse over a link to get help. (b) Short definitions. Here the reader clicks on an underlined term to get a short definition.

design, we began with a simple contrast between learning with the text and learning with HyperCGI. We focused on what participants learned about the semantics of arithmetic word problems and what they learned about the growth and development of children's problem-solving strategies. This contrast between standard text-based materials and HyperCGI is important because hypermedia treatments of existing texts often do not lead to improved learning [29-35]. Because media studies are typically expensive to conduct and often difficult to interpret, we employed a novel single-subject research methodology that afforded sound inference without requiring a large sample. In general, single-subject methods trade fewer participants for greater sampling of each participant's behavior. Hence, the study also provides a forum for assessing the feasibility of employing research designs and methods not typically employed in the field.

Method

Participants

The ten participants (8 female, 2 male) were preservice teacher volunteers from an undergraduate Educational Psychology course. The participants were paid and received class credit.

Design

This study used a modified form of a single-subject, multiple-baseline design called a "regulated randomization procedure" [36]. In contrast to group designs, single-subject designs investigate few individuals over an extended time period, with continual assessment of the dependent variable (e.g., every day for a week). In single-subject designs, the effect of one treatment is contrasted with another treatment (or baseline), but participants are not assigned to treatment groups. Instead, each person participates in all treatments but for variable lengths of time (or beginning at different times, etc.). Measurements of the dependent variables (described later), which are collected throughout the experiment, are used to compare a participant's scores in one phase of the experiment (e.g., a baseline phase) with his or her scores in the next phase of the experiment (e.g., a treatment phase), in order to make statistical conclusions about the efficacy of the treatment.

In multiple-baseline designs (and the regulated randomization procedure used in our study), each participant may spend a different amount of time in the baseline phase, so that the duration of treatment is staggered across subjects [37]. For example, a multiple-baseline design with four participants might specify that one subject (randomly chosen) will switch from baseline to treatment on the third session, another on the sixth session, another on the ninth, and yet another on the twelfth session. The purpose of this staggered intervention is to satisfy a host of critical internal, discriminant, and external validity constraints [37, 38]. Note that in this example design, there are $4! = 24$ possible random assignments of subjects to the four sessions.

The regulated randomization procedure is an extension of the multiple-baseline design that permits more random assignments (and therefore more statistical power), while maintaining the staggered interventions requirements, and thereby key validity constraints. Modifying our earlier four-participant example, a regulated randomization procedure might specify that one participant (randomly chosen) will switch to the treatment phase after the third or fourth session (randomly chosen), another participant will switch to the treatment phase after the sixth or seventh session, another at the ninth or tenth session, and another after the twelfth or thirteenth session (in this example there are $4! \times 2 \times 2 \times 2 \times 2 = 384$ possible assignments). The test statistic compares some measure of the treatment assessments to the baseline assessments. The statistical significance of the experiment is determined by comparing the observed value of the test statistic to the non-parametric distribution of the test statistic (in this distribution there is one value of the test statistic for each possible random assignment).

This study was conducted using ten participants over a period of ten sessions. The baseline in this design was learning about children's mathematical reasoning with traditional texts, and the treatment phase was learning with the HyperCGI system. The nature of these phases is discussed further in the procedure section. Because a complete regulated randomization design consisting of all ten subjects is computationally intractable, the ten participants were assigned to two smaller regulated randomization experiments.[1] Six people participated in the first experiment and four in the second. Both of these smaller experiments were conducted concurrently, differing only in the number of participants enlisted in the design. The results of these two experiments were joined using an additive method for combining probabilities from independent experiments [39]. Hereafter, the results are reported as one (combined) experiment.

Procedure

The ten sessions were conducted every other day (Monday, Wednesday, and Friday). Each session (in both the text and hypermedia phases) consisted of a twenty-minute study period, immediately followed by self-paced assessments of participants' learning that typically lasted five to ten minutes. During the text phase sessions of the experiment, participants studied chapters from CGI materials that are typically provided to pre-service and in-service teachers as an introduction to CGI word-problem types and solution strategies for addition and subtraction (see [27] for a more detailed description). During hypermedia sessions, participants used the hypermedia tool to learn about the same CGI subject matter. In both the text and hypermedia phases, participants were provided with paper and pencils for their own (optional) note taking. Participants were told to study the materials for twenty minutes, in the manner that they best learned. Participant use of HyperCGI ranged from a minimum of two sessions to a maximum of eight sessions.

An extra session was scheduled with each participant between the text and hypermedia phases. This extra session lasted thirty minutes, and was used to train subjects in the use of the hypermedia tool. The first author ran the extra sessions, and only talked about the features and use of the hypermedia system, not about the content or relationships of the CGI materials. The participants were observers

---

[1] Regulated randomization procedures use non-parametric analyses that require that the experimental outcome be computed for each possible randomly assigned text phase duration. A suitable ten participant regulated randomization procedure consists of over fifty-nine billion possible assignments. Such a complex design is not amenable to a calculation given the computational power and storage media of today's microcomputers. The resulting experiments, A and B, consist of 282,240 and 384 possible assignments respectively.

during this extra session, and did not use the software at anytime during this session.

Measures

There are two main sources of data in this experiment: assessments of learning and computer recorded logs of readers' interactions with the hypermedia system.

Assessments of Learning

The assessments were used to determine the relative effectiveness of the text and hypermedia learning environments. The same assessment was administered following each of the sessions and consisted of two computer-administered sorting tasks. In the *word problems* sorting task, the participants saw a computer display of ten arithmetic word problems and an empty four rows by three columns grid. From an expert point of view, the columns of the grid correspond to the status of the unknown quantity in the word problem (i.e., the distinction among $? + 3 = 5$, $3 + ? = 5$, and $3 + 2 = ?$) because the nature of the unknown set has a significant impact on children's thinking. Similarly, from an expert's perspective, the four rows of the grid correspond to the nature of the mathematical action depicted in the word problem (i.e., the distinction among joining two sets, separating two sets, or two different kinds of static comparisons between sets) because these actions also influence how children reason about the problem. Participants sorted the ten problems into the grid with the aim of capturing the underlying semantics of the word problems. The ten word problems were displayed in a different random order on each administration of the assessment.

The second task, the *solution strategies* sort, was a sort of descriptions of twelve different strategies used by children to solve problems. Participants again saw an empty four rows by three columns grid and located solution strategies within this matrix. From an expert perspective, the rows of the grid distinguished among types of problems that typically evoke the associated strategies, whereas the columns corresponded to the developmental level of the strategies (three levels ranging from strategies involving "direct modeling" of the problem context to those employing invented algorithms). The twelve solution strategies were displayed in a different random order on each administration of the assessment.

Scoring

Participants' *word problems* sorts were scored on a 0 to 12 point scale, with higher scores reflecting an organization more consistent with the conceptual model presented in the training materials. Thus, a sort receiving a score of 12 points used one dimension to classify problems into one of the four main problem types of CGI, and the other dimension to rank the problems within a type by the nature of the unknown quantity. For example, consider the word problems in Table 1 (just 5 of the 10 problems used in the sorting task).

The sort that is more consistent with the expert perspective would group problems A, B, and C together in one row of the sorting grid. These are grouped because all three are Join Problems—Problem A is a Join Start Unknown problem ("? + 5 = 11"), problem B is a Join Change Unknown problem ("9 + ? = 14"), and problem C is a Join Result Unknown problem ("15 + 7 = ?"). An expert sort would also rank these problems according to difficulty (i.e., either A, B, C or C, B, A). Sorts that are less consistent with the CGI model might group problems A, D, and E in one row of the sorting grid. Novices are prone to make this grouping because problems A, D, and E all talk about "Darlene" and "Matt," use the numbers "11" and "5," and can be solved with a simple subtraction (11 – 5 = ?).

Participants' *solution strategies* sorts were scored on a 14-point scale. Maximum scores were awarded to participants who coordinated both expert dimensions (developmental level, type of word problem) simultaneously. Intermediate scores were assigned to sorts characterized by one dimension and low scores to more haphazard arrangements (compared to the expert perspective).

Hypermedia Logs

The hypermedia system recorded detailed trails of readers' interactions. The resulting log files provide potentially useful sources of evidence to interpret outcomes in the hypermedia phase of the experiment. The log files contained time-stamped recordings of readers' mouse clicks and key strokes. From these data, it is possible to calculate the sequence of screens that readers saw, the time they spent on each screen, and the tool used to access each screen.

Table 1. Five Word Problems Used in the Sorting Task

A.	B.	C.
Darlene had some stickers. Matt gave her 5 more stickers for her birthday. Now she has 11 stickers. How many stickers did Darlene have to start with?	Tabitha had 9 marbles. Joe gave her some more marbles. Now Tabitha has 14 marbles. How many marbles did Joe give to Tabitha?	Arthur had 15 stuffed animals. John gave him 7 more stuffed animals. How many stuffed animals does Arthur have now?
D.	**E.**	
Darlene has 11 stickers. Matt has 5 stickers. How many more stickers does Darlene have than Matt?	Darlene had 11 stickers. She gave some stickers to Matt. Now she has 5 stickers. How many stickers did Darlene give to Matt?	

RESULTS

For each participant, a series of derived measures were calculated based on the raw scores obtained on the *word problems* and the *solution strategies* sorts. Figure 5 displays the raw scores for one participant on the *word problems* sort task. For this participant, the text phase of the experiment lasted for four sessions, and the hypermedia phase lasted the remaining six sessions. Table 2 displays the both the raw scores and derived measures for the *word problems* sorts. From the *word problems* raw scores, difference scores are calculated for each session (i.e., difference score = session [x] sort score – session [x-1] sort score). These difference scores were used to compare the rates of learning in the text and hypermedia phases. These difference scores were then ranked from smallest to largest to obtain the ranked differences row entries. Ranked differences were used in the statistical analysis of the regulated randomized procedure [36]. The average difference score in each phase (text and hypermedia) is used for descriptive comparisons of the learning rates between phases, while the average ranked difference score is used for statistical comparisons. From these data on Table 2, notice that this participant was gaining an average of 0.6667 points per word problem sort in the text phase (5.0 ranks), compared to an average of 1.500 points in the hypermedia phase (5.0 ranks).

Figure 6 displays learning gains over time (represented as an average difference score) for all ten participants during the text and hypermedia phases. Note that, on average, participants learned more with hypermedia than with text. However, some ceiling effects were noted for the word problem sort, especially for one participant who achieved the maximum possible score by the third session. To remove the effect of this measurement bias, this participant's difference scores were replaced with 0 for all ten sessions on the problem sort task.[2] Thus, this participant was treated as learning at the same rate in both the text phase and hypermedia phase for the problem type measure.

Table 3 summarizes the average difference between sessions in the text and hypermedia phases of the experiment for all participants. Note that participants did not seem to learn much about distinctions among word problems from session-to-session when reading text (an average gain of –0.04). In contrast, their performance increased during each episode of learning with hypermedia (an average gain of 0.50). For solution strategies, inspection of Table 2 indicates that learning about these strategies generally did not improve from session-to-session when reading text (an average gain of –0.40). In contrast, participants' understanding of children's solution strategies typically increased during each episode of learning with hypermedia (an average gain of 0.35). The results of

---

[2] It is not possible to remove this participant entirely from the analyses, primarily because the test is based upon a ten participant randomization scheme. The approach we have used is a conservative one, and favors the hypothesis of no difference between the text and hypermedia materials.

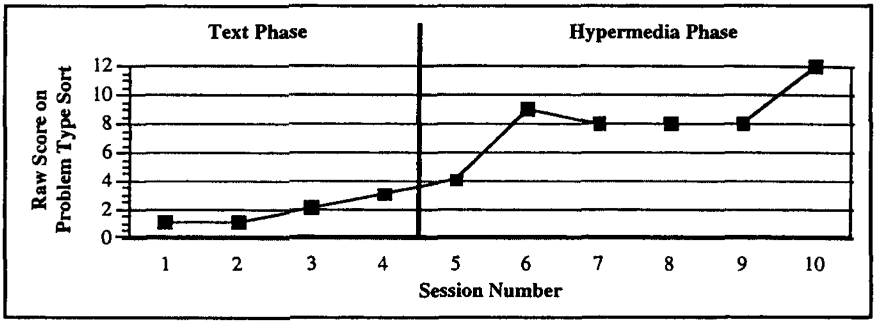


Figure 5. Raw scores for one participant on word problems sorting task.

Table 2. Raw Scores for Word Problems Sorts for One Participant

	Session Number									
	1	2	3	4	5	6	7	8	9	10
Raw scores	1	1	2	3	4	9	8	8	8	12
Difference scores		0	1	1	1	5	−1	0	0	4
Ranked differences		3	6	6	6	9	1	3	3	8
Mean difference		0.6667					1.5000			
Mean ranked difference		5.0000					5.0000			

randomization tests on ranked differences are also provided in Table 3. The learning advantage observed for hypermedia was clearly reliable for the solution strategies and less so for the distinctions among the word problems, perhaps because of the less reliable measurement properties of the word problems sort. Overall, these results suggest that the hypermedia system was clearly an effective improvement compared to the text based materials for learning about children's development of solution strategies, and arguably more effective at conveying the structure of elementary problem types.

Analysis of Log Files

The data in the log files tracked the location and time of each mouse click. From these individual actions, summary measures were constructed. These measures included frequencies of use for each of the navigational features

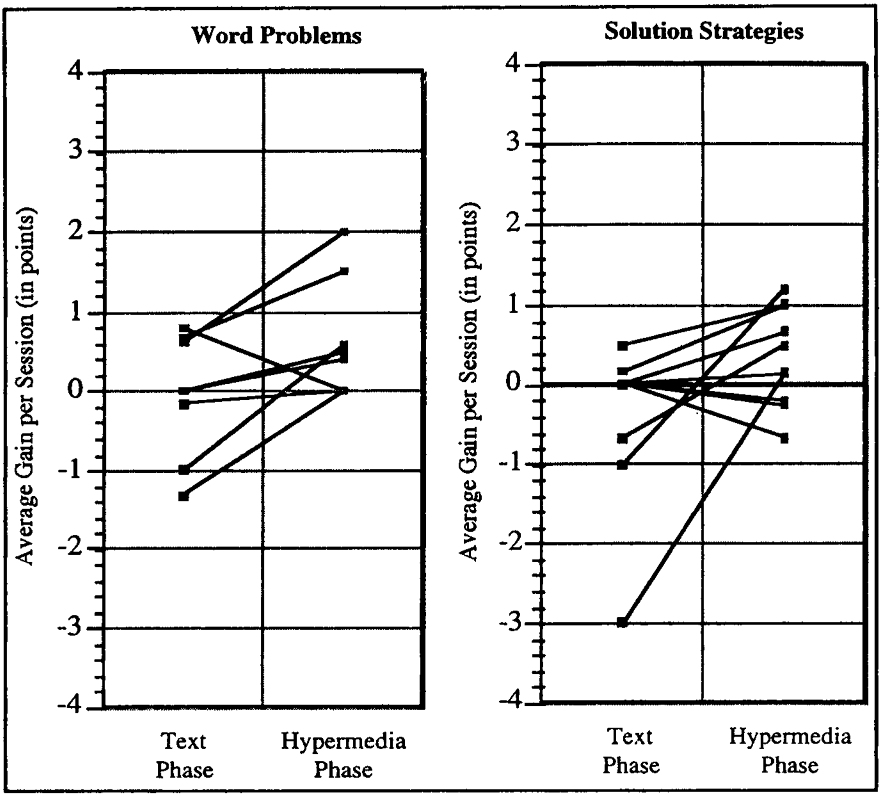


Figure 6. Summary of Results for Measure (Problem Type, Solution Strategy) and Phase (Text, Hypermedia).

Table 3. Statistical Summary of Results by Learning Measure

Learning Measure	Average Difference Score (in Points)		Average Effect (Hypermedia—Text)		
	Text	Hypermedia	(in Points)	(in Ranks)	1-Tailed *p*
Problem Type Sort	–0.04	0.50	0.54	0.73	0.075
Solution Strategy Sort	–0.40	0.35	0.75	1.25	0.009

(Guided Tours, Outline, Index, etc.), as well the time spent viewing each screen, annotation, and video in the system. To investigate the relationship between participants' actions and learning, correlations were computed for each of the measures of activity with the average difference score on each of the two outcome measures (problem type sort, solution strategy sort) in the hypermedia phase. Results from the text phase were not used in any way. Nine of the ten participants used the hypermedia tool for at least two sessions. The tenth participant only used the system for one session, and therefore was not included in this analysis.

Topical Content

The content of each screen was assigned to one of the following: problem type, solution strategy, problem type and solution strategy (some nodes addressed the relationship between problem type and solutions), children's development (some nodes addressed typical developmental trends in the growth of children's solutions), galleries (the example-based learning areas) and navigation (e.g., opening screen, organizers, guided tour endings, etc.). Table 4 summarizes the percentage of their time that participants spent on each topic, and the percentage of these topical screens in the system as a whole. Participants' time on each topic roughly mirrored the amount of the total corpus devoted to that topic. However, participants devoted 12 percent of their learning time to the "problem and solutions" screens, even though these screens represented only 6 percent of the screens in the system. This suggests that participants used these nodes to facilitate "crisscrossing" the conceptual landscape from the perspective of solution strategies and problem types.

Navigational Tools

Each participant used several of the navigational features over the course of the experiment, and several experimented with all the navigational features of the system. The analyses of the guided tours, outline, index, next and prior, return, typed links, bookmarks, and history features all revealed small correlations with

Table 4. Average Percentage of Time Spent on Topics

Topic	Percent of Time	Percent of Screens About this Topic
Problem Types	13	21
Solution Strategies	46	39
Problems and Solutions	12	6
Development	5	12
Galleries	12	6
Navigation	11	15

measures of learning. However, not all navigational tools were used with the same frequency. Across all subjects, the guided tour facility was the most popular, as 42 percent of all screen accesses were reached via the guided tours. The default linear buttons were next with 23 percent, closely followed by the use of the hypermedia links (19%). All other navigational features were used less than 10 percent. However, this overall tool usage pattern was not stable over time. For example, guided tours were popular for the first few sessions on the system and became less popular over time. Conversely, the next, prior, and return features were least popular to start with, but became increasingly more popular with extended usage ($r = .64$ between session and frequency of use). Use of the typed links, outline, index and history features did not appear to fluctuate over the course of this study.

Role of Video Examples

On average, participants spent 43 percent of their time viewing the video examples embedded in the text. The average time that readers spent viewing a video example was a strong predictor for learning gains for the solution strategy sorts ($r = .81, p < .05$) but not for the problem type sorts.

Galleries

In addition to the embedded videos in the text of the hypermedia materials, examples were also provided in the galleries (the examples included in the galleries are not included in the above analyses of video examples). Solution strategy galleries contained video examples, while problem type galleries contained text examples. On average, participants spent 12 percent of their time on the Gallery screens. The percentage of time that individual participants spent using the galleries as a whole strongly predicted learning gains for solution strategies ($r = .70, p < .05$) but not for problem types. However, time spent on the problem type galleries (considered separately from galleries as a whole) more strongly related to gains on the problem type sort ($r = .65$). Time spent on solution galleries (considered separately) were less successful at predicting gains on the solution strategy sort ($r = .45$) than time spent on galleries as a whole. When the role of *embedded video examples* and the *role of the galleries* are jointly considered, a regression analysis indicated that these two factors strongly predicted performance gains on the solution strategy sort; $F(2,6) = 8.03$, $R = .85$, $p < .05$. This suggests a form of "criss crossing." Learning about solution strategies perhaps was embedded in learning about the problem types that tended to elicit them.

DISCUSSION

Learning about the development of children's mathematical reasoning with hypermedia proved significantly better than learning with text. Secondary analyses suggest that this finding can be attributed to several elements of hypermedia design. First, the nonlinear structure of the document enabled learners to "criss-cross" the conceptual landscape. Participants used the criss-cross structure provided by the typed hypermedia links, followed the implicit criss-crossing provided by the Guided Tours, and made extensive use of the portions of the system affording dual viewpoints about problem types and solution strategies. It seems clear that participants were viewing the conceptual landscape of CGI from several perspectives and criss-crossed this landscape quite willingly.

Second, the system made the structure of the domain more visible by employing typed links to signal important relationships in CGI. Readers used these links quite often (19% of the time) throughout the course of this study. The typed linking structure was also visible in the Index tool developed for navigation of the document, so that conceptual structure could serve as navigational beacons.

Third, learning about solution strategies, a relatively ill-structured component of this domain, was related to time engaged with video examples, perhaps because the video clips exemplified aspects of children's strategies not easily conveyed in text. As expected, learning about the structure of word problems was less dependent on examples, perhaps because distinctions among problem types were well-defined and thus more readily conveyed by text.

Fourth, the wide variety of navigational tools provided multiple points and forms of access to learners, so that system use could be aligned with individual differences. The ideal paths provided by the tour facility generally facilitated initial access to the document. Thereafter, readers navigational patterns were characterized more by diversity than prototypicality. For example one participant used a "linear traversal" strategy (a pattern defined by [40]) in the first session, but used a more directed search strategy called "searching" (a pattern defined by [41]) in the next session. Often participants would exhibit two or more navigational patterns in the same session. By providing participants with a wide variety of navigational tools, the interface was flexible enough to support readers' efforts to adapt their navigational pattern to their changing learning goals, expertise, and domain knowledge.

Although it is true that time spent on video examples was predictive of successful learning, it does not follow that the success of learners is solely attributable to access to these video examples. The effects of the videos in learning in HyperCGI was large ($r = .81$) compared to well established research that has shown media effects to be either small [42] or non-existent [43]. The magnitude of the effect size suggests that the learning advantage experienced by the readers of HyperCGI is more than just a difference in format. Rather, because the system was organized around criss-crossing the domain, ideal learning paths, and a learning by example theme, participants were able to access the video examples at very opportune moments (compared to the text materials). That is, because the system was designed very carefully, participants were able to access video examples when they were needed, as many as were needed, and were able to access those examples efficiently.

In addition to lessons learned about hypermedia design, the single-subjects methodology employed in this study could be used profitably in other media-related research. Although we employed a regulated randomization design to support inference about learning, it would be relatively straightforward to make inference about other issues as well. For example, if one wished to test the influence of different suites of tools on navigation in a hyperspace, one could stagger the introduction of first one suite of tools (A) and then another suite of tools (B). One could collect log file data and construct measures of readers' search behavior to make inference about the effects of each suite of tools on search behavior, without the added expense of large numbers of participants. Further investigation of conceptually-based hypermedia design and of the utility of single-subject methodologies both appear warranted.

REFERENCES

1. J. D. Bolter, *The Writing Space*, Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1991.
2. J.-F. Rouet, J. J. Levonen, A. Dillon, and R. J. Spiro, An Introduction to Hypertext and Cognition, in *Hypertext and Cognition*, J.-F. Rouet, J. J. Levonen, A. Dillon, and R. J. Spiro (eds.), Lawrence Erlbaum, Mahwah, New Jersey, pp. 3-8, 1996.
3. J.-F. Rouet and J. J. Levonen, Studying and Learning with Hypertext, in *Hypertext and Cognition*, J.-F. Rouet, J. J. Levonen, A. Dillon, and R. J. Spiro (eds.), Lawrence Erlbaum, Mahwah, New Jersey, pp. 9-23, 1996.
4. C. Chen and R. Rada, Interacting with Hypertext: A Meta-Analysis of Experimental Studies, *Human-Computer Interaction, 11*, pp. 125-156, 1996.
5. S. P. Lajoie and S. J. Derry, *Computers as Cognitive Tools*, Lawrence Erlbaum, Hillsdale, New Jersey, 1993.
6. R. J. Spiro, R. L. Coulson, P. J. Feltovich, and D. K. Anderson, Cognitive Flexibility Theory: Advanced Knowledge Acquisition in Ill-Structured Domains, in *Tenth Annual Conference of the Cognitive Science Society*, V. Patel (ed.), Lawrence Erlbaum, Hillsdale, New Jersey, pp. 375-383, 1988.
7. J. Sweller and P. Chandler, Why Some Material is Difficult to Learn, *Cognition and Instruction, 12*:3, pp. 185-233, 1994.
8. M. T. H. Chi, R. Glaser, and E. Rees, Expertise in Problem Solving, in *Advances in the Psychology of Human Intelligence*, R. Sternberg (ed.), Lawrence Erlbaum, Hillsdale, New Jersey, pp. 7-75, 1982.
9. F. Conklin, Hypertext: An Introduction and a Survey, *Computer*, pp. 17-41, September 1987.
10. M. M. Recker, A Methodology for Analyzing Students' Interactions within Educational Hypertext, *Multimedia and Hypermedia Annual*, 1994.
11. M. A. Just and P. A. Carpenter, A Capacity Theory of Comprehension: Individual Differences in Working Memory, *Psychological Review, 99*, pp. 122-149, 1992.
12. M. A. Britt, J. Rouet, and C. A. Perfetti, Using Hypertext to Study and Reason about Historical Evidence, in *Hypertext and Cognition*, J.-F. Rouet, J. J. Levonen, A. Dillon, and R. J. Spiro (eds.), Lawrence Erlbaum, Mahwah, New Jersey, 1996.

13. M. M. Recker, A. Ram, T. Shikano, G. Li, and J. Stasko, Cognitive Media Types for Multimedia Information Access, *Journal of Educational Multimedia and Hypermedia, 4*:2/3, pp. 183-210, 1995.
14. J-F. Rouet, Interactive Text Processing by Inexperienced (hyper-) Readers, in *Hypertexts: Concepts, Systems, and Applications,* A. Rizk, N. Streitx, and J. André (eds.), Cambridge University Press, Cambridge, England, pp. 250-260, 1990.
15. M. A. Horney and L. Anderson-Inman, The Electrotext Project: Hypertext Reading Patterns of Middle School Students, *Journal of Educational Multimedia and Hypermedia, 3*:1, pp. 71-91, 1994.
16. D. Dee-Lucas, Effects of Overview Structure on Study Strategies and Text Representations for Instructional Hypertext, in *Hypertext and Cognition,* J.-F. Rouet, J. J. Levonen, A. Dillon, and R. J. Spiro (eds.), Lawrence Erlbaum, Mahwah, New Jersey, 1996.
17. L. F. Leader and J. D. Klein, *The Effects of Search Tool and Cognitive Style on Performance in Hypermedia Database Searches,* paper presented at the national convention of the Association for Educational Communications and Technology, ERIC ED 373 729, Nashville, Tennessee, 1994.
18. F. Campagnoni and K. Ehrlich, Information Retrieval Using a Hypertext-Based Help System, *ACM Transactions on Information Systems, 7,* pp. 271-291, 1989.
19. S. Gray, C. Barber, and D. Shasha, Information Search with Dynamic Text vs Paper Text: An Empirical Comparison, *International Journal of Man-Machine Studies, 35,* pp. 575-586, 1991.
20. D. J. Ayersman, Reviewing the Research on Hypermedia-Based Learning, *Journal of Research on Computing in Education, 28*:4, pp. 500-525, 1996.
21. N. Goodman, *Languages of Art: An Approach to a Theory of Symbols,* Hackett Publishing, Indianapolis, Indiana, 1976.
22. S. M. Williams, Putting Case-Based Instruction into Context: Examples from Legal and Medical Education, *Journal of the Learning Sciences, 2*:4, pp. 367-427, 1992.
23. M. Ward and J. Sweller, Structured Working Examples, *Cognitive and Instruction, 8,* pp. 1-39, 1990.
24. A. Black, P. Wright, D. Black, and K. Norman, Consulting On-Line Dictionary Information while Reading, *Hypermedia, 4*:3, pp. 145-169, 1992.
25. B. Shneiderman, *Designing the User Interface: Strategies for Effective Human-Computer Interaction,* Addison-Wesley Publishing, Reading, Massachusetts, 1992.
26. P. G. Polson and C. H. Lewis, Theory-Based Design for Easily Learned Interfaces, *Human Computer Interaction, 5,* pp. 191-220, 1990.
27. T. P. Carpenter, E. Fennema, P. L. Peterson, C. Chiang, and M. Loef, Using Knowledge of Children's Mathematics Thinking in Classroom Teaching: An Experimental Study, *American Educational Research Journal, 26,* pp. 499-531, 1989.
28. T. Mayes, M. Kibby, and T. Anderson, Learning about Learning from Hypertext, in *Designing Hypermedia for Learning,* D. Jonasen and H. Mandl (eds.), Springer Verlag, Berlin, pp. 227-250, 1990.
29. J. W. Rojewski, J. P. Gilbert, and C. A. Hoy, Effects of a Hypertext Computer Program on Academic Performance in a Special Education Course for Nonmajors, *Teacher Education and Special Education, 17*:4, pp. 249-259, 1994.

30. R. Azevedo, S. G. Shaw, and P. M. Bret, *The Effectiveness of Computer-Based Hypermedia Teaching Modules for Radiology Students*, paper presented at the annual meeting of the American Educational Research Association, ERIC ED 385 187, San Francisco, California, April 1995.
31. C. McKnight, A. Dillon, and J. Richardson, A Comparison of Linear and Hypertext Formats in Information Retrieval, in *Hypertext: States of the Art*, R. McAleese and C. Greene (eds.), Intellect, Oxford, England, 1990.
32. R. Aust, M. Kelley, and W. Roby, The Use of Hyper-Reference and Conventional Dictionaries, *Educational Technology Research and Development, 41*:4, pp. 63-73, 1993.
33. L. P. McCoy, Decisions in Inferential Statistics with HyperCard: Design and Field Test, *Computers in the Schools, 10*:1-2, pp. 69-77, 1994.
34. D. A. Becker and M. M. Dwyer, Using Hypermedia to Provide Learner Control, *Journal of Educational Multimedia and Hypermedia, 3*:2, pp. 147-162, 1994.
35. D. Egan, J. Remde, L. Gomez, T. Landauer, J. Eberhardt, and C. Lochbaum, Informative Design-Evaluation of SuperBook, *ACM Transactions on Information Systems, 7*, pp. 30-57, 1989.
36. M. J. Koehler and J. R. Levin, Regulated Randomization: A Potentially Sharper Analytic Tool for the Multiple-Baseline Design, *Psychological Methods, 3*:2, pp. 206-217, 1998.
37. A. E. Kazdin, *Research Design in Clinical Psychology* (2nd Edition), Allyn and Bacon, Boston, 1992.
38. J. R. Levin, On Research in Classrooms, *Mid-Western Educational Researcher, 5*:2, pp. 2-16, 1992.
39. E. S. Edgington, Randomization Tests for One-Subject Operant Experiments, *Journal of Psychology, 90*, pp. 57-68, 1972.
40. M. A. Horney, Case Studies of Navigational Patterns in Constructive Hypertext, *Computers & Education, 20*:3, pp. 257-270, 1993.
41. D. Canter, R. Rivers, and G. Stores, Characterizing User Navigation through Complex Data Structures, *Behaviour and Information Technology, 4*:2, pp. 93-102, 1985.
42. R. B. Kozma, Learning with Media, *Review of Educational Research, 61*:2, pp. 179-211, 1991.
43. R. Clark, Reconsidering Research on Learning from Media, *Review of Educational Research, 53*, pp. 445-459, 1983.

Direct reprint requests to:

Matthew J. Koehler or Richard Lehrer
Department of Educational Psychology
1025 W. Johnson St.
University of Wisconsin
Madison, WI 53706