Review of Educational Research. Summer 1991, Vol. 61, No. 2, pp. 179-211

Learning With Media

Robert B. Kozma
University of Michigan

This article describes learning with media as a complementary process within which representations are constructed and procedures performed, sometimes by the learner and sometimes by the medium. It reviews research on learning with books, television, computers, and multimedia environments. These media are distinguished by cognitively relevant characteristics of their technologies, symbol systems, and processing capabilities. Studies are examined that illustrate how these characteristics, and the instructional designs that employ them, interact with learner and task characteristics to influence the structure of mental representations and cognitive processes. Of specific interest is the effect of media characteristics on the structure, formation, and modification of mental models. Implications for research and practice are discussed.

Do media influence learning? The research reviewed in this article suggests that capabilities of a particular medium, in conjunction with methods that take advantage of these capabilities, interact with and influence the ways learners represent and process information and may result in more or different learning when one medium is compared to another for certain learners and tasks.

This article responds to a challenge by Clark (1983) for "… researchers [to] refrain from producing additional studies exploring the relationship between media and learning unless a novel theory is suggested" (p. 457). He extended this challenge after reviewing the existing comparative research on media and concluding that "… media do not influence learning under any conditions" (p. 445).. Rather, "... media are mere vehicles that del' instruction but do not influence student achievement any more than the truck that delivers our groceries causes changes in our nutrition" (p. 445). The theoretical framework supported by the review herein presents an image of the learner actively collaborating with the medium to construct knowledge. It stands in vivid contrast to an image in which learning occurs as the result of instruction being "delivered" by some (or any) medium. The framework is meant to provide the novel approach required by Clark before research on media and learning can progress.

In this theoretical framework, learning is viewed as an active, constructive process whereby the learner strategically manages the available cognitive resources to create new knowledge by extracting information from the environment and integrating it

This project was performed pursuant to grant number OERI-GO086-90010 from the Office of Educational Research and Improvement/Department of Education (OERI/ED). The opinions expressed herein do not necessarily reflect the position or policy of the OERI/ED or the Regents of The University of Michigan, and no official endorsement by either group should be inferred. The author acknowledges the comments of many colleagues on earlier drafts, particularly those of Patricia Baggett, Jerome Johnston, Wilbert McKeachie, Paul Pintrich, Nancy Nelson Spivey, and an anonymous reviewer. Also acknowledged is support from the Dana Foundation Fellowship for Computer-Assisted Instruction in the Humanities at Carnegie Mellon University.

with information already stored in memory. This process is constrained by such cognitive factors as the duration and amount of information in short-term memory, the task-relevant information that is available in long-term memory, the structure of this information, the procedures that are activated to operate on it, and so on. Consequently, the process is sensitive to characteristics of the external environment, such as the availability of specific information at a given moment, the duration of that availability, the way the information is structured, and the ease with which it can be searched.

The relationship between the internal and external cognitive environments is explicitly addressed by the emerging discussion of *distributed cognition.* There are two perspectives one can take in this discussion: a system view or a personal view (Norman, 1989). Pea (1990) and Perkins (1990) take a system perspective and examine how cognition within the system is augmented by its distribution among individuals and between individuals and artifacts (e.g., computers, calculators, etc.) The theoretical framework developed in this review approaches distributed cognition from the perspective of the individual. This review examines the effects that the sharing of cognition between an individual and a medium has on the cognitive representations and processes of that individual, particularly those effects that endure beyond the immediate interaction (Salomon, 1990).

The subdomain of the external environment examined in this article is *mediated information*, not only that information that is intentionally educational (e.g., a computer-based lesson) but also other information that may not have an explicit educational goal (e.g., in popular books, television programs, etc.). This review does not directly address information embedded in what are sometimes called *authentic situations* (Brown, Collins, & Duguid, 1989), but it complements learning in such situations. Nor does this article examine the larger social environment within which mediated interactions occur (Perkins, 1985). Although it may be the above contexts, and the ways media are integrated into them, that have the greatest impact on how people think and learn, these broader contexts will only be referenced here. The primary focus of this article is finer grained. It will examine the specific episodes within which a learner interacts with mediated information to influence learning.

This article will provide a definition of media and use it to examine the theoretical and research literature on learning from books, television, computers, and multimedia environments. Each section will examine how the complementary construction of representations, and operations performed on them, is influenced by characteristics of the medium, designs that take advantage of these characteristics, and the characteristics of learners and tasks. The intent is to demonstrate the relative cognitive effects of learning with different media, particularly effects related to the structure, formation, and modification of mental models.

Media Defined

Media can be defined by its technology, symbol systems, and processing capabilities. The most obvious characteristic of a medium is its technology: the mechanical and electronic aspects that determine its function and, to some extent, its shape and other physical features. These are the characteristics that are commonly used to classify a medium such as a television, a radio, and so on. The cognitive effects of these characteristics, if any, are usually indirect. Characteristics such as size, shape, and weight make it more likely that a student will learn with a book but not a

computer while on a bus, although of course this predilection is changing as computers get smaller, fighter, and cheaper. A few cognitive effects of technology, however, are more direct. For example, the size and resolution of many computer screens are such that reading their texts may be more difficult than reading the text of some books (Haas, 1989).

However, the primary effect of a medium's technology is to enable and constrain its other two capabilities: the symbol systems it can employ and the processes that can be performed with it. For example, a computer with a graphics board or a speech synthesis board can use different symbols in its presentations than those without these features. Computers with enough memory to run expert systems can process information in different ways than those without such a memory. These additional symbol systems and processes are likely to account for the cognitive effects of these systems, rather than the technology, per se.

Symbol systems and processing capabilities have a number of important implications for learning. Salomon (1974, 1979) describes the relationship between a medium's symbol systems and mental representations. Symbol systems are *modes of appearance* (Goodman, 1976), or sets of elements (words, picture components, etc.) I that are interrelated within each system by syntax and are used in specifiable ways in relation to fields of reference. (Words and sentences in a text may represent people, objects, and activities and be structured in a way that forms a story.) A medium can be described and perhaps distinguished from other media by its capabilities to employ certain symbol systems. Thus, television can be thought of as a medium that is capable of employing representational (i.e., pictorial) and audio-linguistic symbol systems (among others). Such characterizations can also be used to specify a certain overlap or equivalence of media. Thus video and motion film can be thought of as equivalent in this regard, while they can be distinguished from radio which can employ only a subset of these symbol systems.

Salomon (1974, 1979) suggests that these characteristics should be used to define, distinguish, and analyze media because they are relevant to the way learners represent and process information from a medium. He contends that certain symbol systems may be better at representing certain tasks and that information presented in different symbol systems may be represented differently in memory and may require different mental skills to process. The research reviewed here supports and elaborates on this contention. For example, studies will be examined that illustrate how symbol systems characteristic of certain media can connect mental representations to the real world in a way that learners with little prior knowledge have trouble doing on their own without the representation of information in these symbol systems.

But, as will be demonstrated, symbol systems alone are not sufficient to describe a medium and its cognitive effects. Information is not only represented in memory; it is processed. Media can also be described and distinguished by characteristic capabilities that can be used to process or operate on the available symbol systems. Thus, information can be searched or its pace of progression changed with videodisc in a, way that is not possible with broadcast video. Including processing attributes in the definition of media can create useful distinctions between videodisc and broadcast video, even though both have access to the same symbol systems. Computers are, of course, especially distinguished by their extensive processing capabilities rather than their access to a particularly unique set of symbol systems.

The processing capabilities of a medium can complement those of the learner; they may facilitate operations the learner is capable of performing or perform those that

the learner cannot. As Salomon (1988) points out, if such processes are explicit and fall within what Vygotsky (1978) calls the *zone of proximal development*, the learner may come to incorporate them into his or her own repertoire of cognitive processes. This review will examine research that illustrates how the processing capabilities of certain media modify and refine the dynamic properties of learners' mental models.

However, it is important to remember that whereas a medium can be defined and distinguished by a characteristic cluster, or profile, of symbol systems and processing capabilities some of these capabilities may not be used in a particular learning episode (Salomon & Clark, 1977). For example, a particular video presentation may use few or no representational symbols (e.g., a talking head presentation). Or, a viewer may allow a videodisc presentation to play straight through and not use the available search capabilities. In these cases, a *virtual medium* is created that consists of the profile of symbol systems and processing capabilities that were actually used during the session: In effect, a television becomes a radio; a videodisc player becomes a broadcast television. It is only the capabilities of the virtual medium that can be expected to have an effect on learning processes and outcomes.

Whether or not a medium's capabilities make a difference in learning depends on how they correspond to the particular learning situation-the tasks and learners involved-and the way the medium's capabilities are used by the instructional design. Tasks vary in their situational characteristics and in the demands they place on the learner to create mental representations of certain information and to operate on that information in certain ways. Learners vary in their processing capabilities, the information and procedures that they have stored in long-term memory, their motivations and purposes for learning, and their meta-cognitive knowledge of when and how to use these procedures and information.

Many learners, perhaps most, can and frequently do supply useful representations and operations for themselves from the information externally available, regardless of the medium used. But learners will benefit most from the use of a particular medium with certain capabilities (as compared to the use of a medium without these) if the capabilities are employed by the instructional method to provide certain representations or perform or model certain cognitive operations that are salient to the task and situation and that the learners cannot or do not perform or provide for themselves. These representations and operations, in turn, influence problem solving and the ability to generate and use representations in subsequently encountered situations. This view of learning with media as a continuous, reciprocal interaction between person and situation-between learner and mediated information-is compatible with Snow's (1989) evolving aptitude-treatment interaction theory.

Learning With Books

The most common medium encountered in school learning is the book. As a medium, books can be characterized by the symbol systems they can employ: text and pictures. The following sections of the review will examine the cognitive processes used in processing text and in processing text with pictures. They will discuss how a distinctive characteristic f this technology-its stability-influences the processing of these symbol systems to construct knowledge representations and how these, in turn, are influenced by the individual differences of learners, primarily differences in their prior domain knowledge. The summary will describe how these processes and structures can be supported by the author when designing a book.

The reading processes and the stability of the printed page. The primary symbol system used in books consists of orthographic symbols that, in Western culture, are composed of phonemic graphemes, horizontally arrayed from left to right. at this arrangement is stable distinguishes text in books from other technologies that use the same symbol system-for example, the marquee on Times Square. This stability also has important implications for how learners process information from books. Specifically, the stability of text aids in constructing a meaning of the text. Learning with text involves the construction of two interconnected mental representations: a *textbase* and a *situation model* (Kintsch, 1988, 1989). The *textbase* is a mental representation derived directly from the text, at the level of both micro- and macrostructure; it is a *propositional* representation of the meaning of the text. While progressing through the text, the reader assembles propositions and integrates them witth those previously constructed. As memory limits are reached, the most recent frequently encountered propositions are retained in short-term memory and Id together by repetition or the embedding of arguments (Kintsch & van Dijk, 1978). The reader generalizes from these local propositions to form *macropropositions,* or summary-like statements that represent the gist of the text. Integrating the from the text in this way increases the likelihood that it will survive in memory and be fixed in long-term memory.

The situation model is a mental representation of the situation described by the text , 1988, 1989). Whereas the textbase is propositional, the situation model can constructed from propositions or spatial information. The situation model is connected to and constructed from information in the text base and from knowledge structures evoked from long-term memory by information appearing early in the text information activated by the reader's purpose. These structures – called, variously, *schemata* (Anderson, Spiro, & Anderson, 1978), *frames* (Minsky, 1975), and *scripts* (Schank & Abelson, 1977) - can be characterized as a framework with a set of labeled slots in which values are inserted for particular situations. These structures serve two related purposes: They provide a scaffold upon which the situation model is from the textbase, and they provide default values so that the reader can inferences about the local situation that were not explicitly mentioned in the Learning from text involves the integration of these representations into the Is knowledge system by updating the schemata currently in long-term memory or by constructing a new schema for an unfamiliar situation.

But, what does any of this have to do with media? How does this symbol system influence mental representations and cognitive processes in distinctive ways? And would learning processes and outcomes be any different for books-which store symbols in a fixed, stable way-then they would be for another medium, say audiotape or lecture, which may convey the same linguistic information but a different symbol system and in a transient way (i.e., speech)?

In many situations for fluent readers, reading progresses along the text in a forward at a regular rate, and the information could just as well be presented in another, more transient medium. But, on occasion, reading processes interact with knowledge and skill in a way that relies heavily on the stability of text to aid and learning.

In the obvious case, the effort required of poor readers to decode the text draws on cognitive resources that would otherwise be used for comprehension, thus increasing the risk of comprehension, or learning, failure (LaBerge & Samuels, 1974). But even

fluent readers may have difficulty with longer or novel words, such as technical terms in an unfamiliar domain. In both of these cases, readers will use the stability of text to recover from comprehension failure. When encountering difficulty, readers win slow their rate, making more or longer eye fixations (Just & Carpenter, 1987), or they may regress their eyes, going back to review a word as an aid to retrieving a meaning for it from memory (Bayle, 1942). Alternatively, readers may retrieve several meanings for a word and may make longer or additional fixations or may regress over a phrase, a clause, or even a sentence to determine which is appropriate for a given context (Just & Carpenter, 1987; Bayle, 1942). Such difficulties might arise from unusual syntactic structures (e.g., The thief stood before the black-robed judge entered the courtroom.) or difficulties in interpreting combinations of words to construct local propositions. Readers will slow their rate for a passage on a difficult or novel topic (Buswell, 1937) when they encounter information within a passage that is particularly important to the meaning of the text (Shebilske & Fisher, 1983) or when they must integrate less well organized sentences into macropropositions (Shebilske & Reid, 1979).

All of these are examples of how readers use the stability of the symbol system in books to slow their rate of progression or even to regress over text in a way that would seem difficult or impossible to do with audiotape's ever-advancing presentation of information. However, this distinction is likely to be crucial only in certain situations. For example, readers in the Shebilske and Reid study (1979) reduced their rate from 302 words per minute to 286. This difference is statistically significant, but it may not have practical significance with regard to media use because the typical audiotape presentation rate of 110-120 words per minute would seem to be slow enough to accommodate

these comprehension difficulties. Even the apparent inability to regress over speech might be accommodated by the 2-second duration of information in acoustic memory (Baddeley, 1981) that would allow a listener to recover the three or four most recently spoken words and achieve the same effect as regression over text. The clearest advantage to the use of the stability of text to aid comprehension is when the reader must regress over segments of information larger than a phrase.

Perhaps more important than the use of the stability of text to recover from local comprehension failure in novel or difficult situations is this use in conjunction with highly developed reading skills (such as those described by Brown, 1980) and elaborate memory structures to strategically process large amounts of text within very familiar domains. This is most dramatically illustrated in a study by Bazerman (1985), who interviewed seven professional physicists and observed them reading professional material in their field. These readers read very selectively, making decisions based on highly developed schemata that extended beyond extensive knowledge of accepted facts and theories in the field to include knowledge about the current state of the discipline and projections of its future development as well as personal knowledge and judgments about the work of colleagues. Readers used this specialized, or *domain,* knowledge to serve their reading purposes. Most often their interests were to find information that might contribute to their immediate research goals or to expand their background knowledge of the field, and they made their selections based on these purposes.

Bringing schemata and purposes to bear, these subjects would typically read by scanning rapidly over tables of contents and by using certain words to trigger their attention to question a particular title more actively. If a particular term attracted their attention, they would look at other words in the title with the result that about

---

p 184

two thirds of the titles more closely examined were subsequently rejected based on this additional information. If even more information was needed to make further selections, they would turn to the abstract.

Having identified an article of interest, they would read parts of it selectively an non-sequentially, jumping back and forth, perhaps reading conclusions then introductions, perhaps scanning figures, and finally reading those sections more carefully that fit their purpose. If an article did not readily fit with their comprehension schemata, the readers would weigh the cost of working through the difficulty against the potential gain relative to their purposes. If they chose to read through a difficult article or section, they would occasionally pause at length to work through the implications of what had been read or read it through several times. They might also ok up background material in reference works and textbooks.

The studies above show the range of ways that readers take advantage of the stable of text to aid comprehension. In the Bazerman (1985) study, strategic readers with considerable domain knowledge would sometimes progress through the at a rapid rate, using a single word to skip a vast amount of information. Other times, they would slow considerably, moving back and forth within a text and across , to add to their understanding of the field. In other studies (Bayle, 1942; Shebilske & Reid, 1979), readers encountering difficulties with unfamiliar words, syntactic structures, or ideas used the stability of the printed page to slow their rate regress over passages. None of these processing strategies are available with the transient, linguistic information presented in audiotape or lecture.

Multiple symbol systems: Learning with text and pictures. Orthographic symbols , of course, not the only ones available to books. Pictures and diagrams are used in books from primers to college textbooks to technical manuals. But, how do readers pictures? What is the cognitive effect of pictures in combination with text? And, does the stability of these symbols, as presented in books, influence this process compared to another medium - say, television - that presents linguistic and pictorial symbols in a transient way? The following section examines the cognitive effects of s and text. The subsequent section directly addresses learning with television. There is a large body of comparative research on learning from text with and pictures. Almost all of the studies examine only the impact on cued recall and traditional experimental designs of the type criticized by Clark (1983). However, there is a consensus, among the reviews of this research, that pictures have positive effects under certain conditions. Pressley (1977), Schallert (1980), and Levie and (1982) generally concur that the use of pictures with text increases recall, for poor readers, if the pictures illustrate information central to the text, en they represent new content that is important to the overall message, or when they depict structural relationships mentioned in the text. The problem with this type research is that it does not reveal the mechanism by which pictures and text influence the learning process.

The four studies below examine processes of comprehension and learning with text pictures. In brief, it appears that the use of both symbol systems facilitates the of the textbase and the mapping of it onto the mental model of the situation. This is particularly facilitating for learners who have little prior knowledge the domain.

A study by Rusted and Coltheart (1979) examined the way good and poor fourth-grade readers used pictured text to learn about physical features, behavior, and

---

habitat of unfamiliar animals. Including pictures of animals in their environments along with the text resulted in better retention by both good and poor readers over the use of text alone. It facilitated retention of all information by good readers but only pictured information (i.e., recall of physical features) by poor readers. Observations of good readers showed that they spent time initially looking at the pictures and rarely looked at them once they started reading. Poor readers, on the other hand, frequently moved back and forth between text and pictures. The process data was not detailed enough to be definitive, but it suggested that good readers used the pictures to evoke an animal schema that guided their reading and aided their comprehension. Poor readers moved back and forth, perhaps, to facilitate the decoding of particular words and to aid in building a mental model of these unfamiliar animals and their habitats.

Stone and Glock (1981) obtained similar findings, using more precise measures, when they examined the reading of second- and third-year college students. Subjects used either text without pictures or pictured text to learn how to assemble a toy pushcart. The text-only group made significantly more assembly errors, particularly errors of orientation. The pictured-text group was most accurate in its constructions, making only 18% of the errors of the text-only group. Eye-tracking data indicated two patterns of picture use. Readers would typically spend the first few seconds examining the picture. Then they would look from text to picture as they progressed through the passage, spending an average of more than 80% of their time looking at text rather than pictures. As in the Rusted and Coltheart study (1979), the data suggest that readers initially use the pictures to evoke a schema that serves as a preliminary mental model of the situation. Subsequently, it seems that the text carries the primary semantic message while the pictures are used to map this information on to this preliminary mental model, elaborating on the components of the push cart and their relative arrangement.

The usefulness of pictures seems to interact with domain knowledge. In a study by Hegarty and Just (1989), college students were tested on mechanical ability and assigned to either a short text or a long text describing a pulley system. The short text merely named the components of the system and described how it operated. The long text also elaborated on the arrangement and structure of the components in the system. All texts were accompanied by a schematic diagram of the pulley system. Precise eye-fixations measured the number and duration of movements back and forth between particular words in the text and specific locations in the diagram. There was a non-significant interaction such that low ability students spent more time than high ability students looking at the diagram when it accompanied the longer text which described the relationship among the components of the system. The high ability students spent more time examining the diagram with the shorter text. The results suggest that. people low in mechanical ability have difficulty forming mental models of mechanical systems from text and use diagrams to help them construct this representation. People with high mechanical ability seem to construct this model from prior knowledge and information from the text, without need to refer to the picture. Interestingly, these high ability students are better able to encode new information from a diagram when the text does not describe all the information relevant to understanding a mechanical system.

In a study by Kuntz, Drewniak, & Schott (1989), university students majoring in either geography or social science read passages that contained concepts and rules on

meteorology. They received text either with or without two types of supplements: (a) representational pictures depicting spatial arrangement, appearance, and configuration of clouds and (b) a tree diagram, that provided an overview of the main concepts, constituting the macrostructure of the text. Students were divided on prior domain knowledge. For students with a higher prior knowledge, the examination of representational pictures did not correlate with posttask comprehension, and the use of the tree diagram correlated negatively with performance. In contrast, subjects with low prior knowledge did better if they both inspected the representational picture very often and spent some time examining the tree diagram. These data suggest that students with little prior knowledge benefited most from the pictures and the tree diagram. Students with sufficient prior domain knowledge relied instead on their own, well-developed mental models to aid comprehension. Indeed, the tree diagram may have conflicted with the idiosyncratic structure of these students' domain knowledge and actually interfered with their comprehension.

These studies may also explain Pressley's conclusion (1977) in his review of studies of text and imaging. He found that learners who do not receive pictures but are instructed to generate images during the processing of story prose recall as much as those who receive pictures and more than those who do not receive pictures and are not instructed

to generate images. However, there were developmental differences. Children of 8 years and older could gainfully generate and use images during text processing, whereas those under the age of 6 appeared unable to generate useful images in response to text, even when directed to do so. In these studies, age may be a surrogate measure for accumulated world knowledge that allows older children to generate mental models that supplement the text and aid comprehension and recall. Younger children may not have sufficient world knowledge to generate such mental models. Thus, they benefit most from pictures to aid this process.

Greeno (1989) elaborates on the situation model in a way that can be useful in analyzing the relationship between text, pictures, cognitive structures, and processes. Greeno proposes a theoretical framework that defines knowledge as a relationship between an individual and a social or physical situation rather than as a property of an individual only. This framework extends the information processing paradigm, that focuses primarily on internal structures and process, to include structures and processes external to the learner. This relativistic notion of knowledge depends heavily on a model of the situation and has considerable implications for learning with media.

In the framework, objects and events organized in relation to human activities (e.g., hitting a ball, buying and selling merchandise), as well as related abstractions (e.g., force, profit margin), are expressed within our culture in various symbolic notations and structures (verbal descriptions, diagrams, graphs, etc.). Mental representations, or mental models, are derived from these symbolic notations and structures and correspond to real world objects and events and their abstractions. [1] These mental models consist of symbolic objects, or mental entities, that may have properties associated with the symbol systems from which they were derived (e.g., arrows representing force vectors) as well as properties of objects in situations that the symbolic structures represent (e.g., balls moving through space and time along certain trajectories). Greeno contends that people can reason in this mental space to solve problems by operating on these symbolic objects in ways that correspond to operations in real situations.

However, too often in school learning, these mental objects and operations have little correspondence to real world objects, events, and their abstractions and map only onto the symbolic domains from which they were derived. The research above suggests that, for some learners, the use of pictures, in addition to text, may provide information needed to map mental representations derived from the text onto mental representations of the real world. This may be due to the fact that pictorial symbol systems share more properties with the corresponding objects and events in the real world than do linguistic symbol systems.

Summary and implications. We now have a picture of learning with books that illustrates the relationship between human information processes and the characteristic stability and symbol systems of the medium. Readers move along a line of text constructing a representation of the textbase. They build a mental model of the situation described with information from the textbase and schemata activated in long-term memory. They slow down to comprehend difficult or important points, and stop or regress to retrieve the meaning of an unfamiliar word or a confusing clause or sentence. They may also use their knowledge of the domain and highly developed strategies to read very selectively in service of a particular purpose they bring to the task. They use titles and abstracts to skip sections or entire articles or to focus on sections of interest. They read summaries, then overviews, reread portions, and move back and forth between texts.

If a picture is available, they may refer to it to supplement the text. An initial look at the picture will evoke domain knowledge, for those that have it. In a less familiar domain, readers will move back and forth frequently between text and picture to clarify the meaning of a word or to construct or to elaborate on a model of the situation. All of these strategies and their resulting mental representations are influenced by the knowledge and purpose the reader brings to the task, by the symbol systems, and by the stability of code that characterizes the book.

An author can use these capabilities in a way that complements the learner's skills and deficiencies. Authors can use the stability of text and pictures in books and knowledge of comprehension processes to design structures within their books that support and facilitate learning. Such structures may include titles (Brandsford & Johnson, 1972), postquestions (Wixson, 1984), explicitly stated behavioral objectives (Mayer, 1984), cohesive text elements (Halliday & Hasan, 1976), signals (Meyer, 1975, 1985; Mayer 1984), and so on. For example, in the Brandsford and Johnson (1972) study, one group of students had considerable difficulty comprehending a paragraph even though it was linguistically simple and contained no difficult words, constructions, or complex concepts. A second group was presented the same paragraph, but this time the paragraph was preceded by a title. In this second condition, the subjects rated the paragraph as more comprehensible, and they recalled it better. Presumably, the title evoked an appropriate schema that allowed the readers to supply information not explicit in the paragraph but important for its comprehension. Other text strategies might evoke different reading processes, such as conducting backward reviews to facilitate retention (Wixson, 1984), focusing attention on certain types of information, or building internal

connections among concepts in the text (Mayer, 1984). Such devices designed into the text can support the purpose and schema-driven strategies evident in the Bazerman study (1985), at least for students with sufficient prior knowledge.

An understanding of the cognitive function of pictures can also inform instructional practice. This understanding can provide text authors with information that

can be applied heuristically to identify situations where pictures would be useful and to design pictures which would accommodate particular learners and tasks (Winn, 1989). Such guidelines may suggest the positioning of pictures in the text, the degree of realism, and the use of arrows and other highlighting mechanisms. For example, the research above suggests that for knowledgeable readers, pictures should be placed early in the text if they are used at all. On the other hand, a less knowledgeable readership would benefit from interspersed pictures, juxtaposed with the corresponding text. Winn (1989) reviews research that suggests that the use of arrows in pictures to highlight critical attributes of objects can facilitate subsequent identification but that the inclusion of details in an illustration can actually interfere with the learning of an object's structure or function.

Learning With Television

Television differs in several ways from books that may affect cognitive structures and processes. As with books, television can employ pictures, diagrams, and other representational symbol systems, but, in TV, these symbols are transient and able to depict motion. Linguistic information in television can be orthographic, but more often it is oral and, as with audiotape and radio, transient. Because in television linguistic and pictorial symbol systems are transient and because they are presented simultaneously, viewers may process this information in a very different way than the back-and-forth serial processing of linguistic and representational information in books. It is also possible that the symbol systems used and their transient nature affects the mental representations created with television.

Television's window of cognitive engagement. Popular notions of TV viewing portray children as staring zombie-like at the screen, but reality is much different. When alternative activities are available, children generally look at and away from the TV between one and two hundred times an hour (Anderson & Field, 1983). Visual attention increases from very low levels during infancy to a maximum during the late elementary school years, declining somewhat during adulthood (Anderson, Lorch, Field, Collins, & Nathan, 1986). Although the median look duration is usually only several seconds, extended episodes as long as a minute are not rare. Looks as long as ten minutes are exceptional. This discontinuous, periodic attention to a medium whose information streams by ceaselessly has important implications for comprehension and learning.

Research indicates that visual attention is influenced by several factors. One set of factors, termed *formal features* by Huston and Wright (1983), includes the use of different types of voices (e.g., children, adult male, adult female), laughing, sound effects, music of different types, animation, cuts, zooms, pans, and so forth. Children's moment-to-moment visual attention may wander from the set, but evidence suggests that they continually monitor the presentation at a superficial level, such that their visual attention is recaptured by certain audio cues. Features that are associated with the onset of visual attention are women's and children's voices, laughter, peculiar voices, sound effects, auditory changes, and visual movement (Anderson, Alwitt, Lorch, & Levin, 1979). Features associated with continued viewing are special visual effects, pans, and high physical activity. The offset of visual attention among children frequently corresponds to the use of men's voices, long zooms, and inactivity.

This image of visual attention seems bottom-up and data-driven, but other evidence suggests that these formal features come to be seen by children as correspond-

ing to the presentation of more or less meaningful content, and it is this second factor, the meaningfulness or comprehensibility of the presentation, that guides visual attention. For example, Anderson, Lorch, Field, and Sanders (1981) found that visual attention to segments of *Sesame Street* was greater for normal segments than for the same visual presentation for which comprehensibility was experimentally reduced by using backward speech or a foreign language. Anderson and Lorch (1983) hypothesize that, through extensive viewing experience, children come to acquire knowledge about the associations between the typical use of various formal features and the

likelihood that the corresponding content will be meaningful and interesting. For example, men's voices may be perceived as generally corresponding to adult-oriented content that is less comprehensible and less interesting to children, and thus male voices do not recruit their visual attention.

Huston and Wright contend that this comprehensibility influences attention in an inverted-U relationship. Content that is very simple or very difficult to comprehend maintains attention less well than content in an intermediate range of difficulty. This creates a window of cognitive engagement, one that is perhaps different for each viewer. Yet, within this window, Huston and Wright (1983) conclude that visual attention is necessary though not sufficient for comprehension; even with visual attention, the depth of comprehension varies.

Salomon (1983) introduces the construct of *amount of invested mental effort*, or AIME, to account for the difference between what is viewed and the depth of comprehension. AIME distinguishes the deep, effortful, nonautomatic elaboration of encountered material from the mindless or shallow processing of information that results in less learning. AIME is in turn influenced by several factors: One is the attitudes people have about the amount of effort required to process a medium's messages; the other is the purpose that people bring to the task.

Salomon (1984) found that a sample of sixth graders rated TV as an easier medium from which to learn than books. When assigned to view comparable stories from television or print, the effort spent on learning reported by the reading group was significantly greater than that reported by the group that viewed the television program. Both groups scored the same on a test of factual recognition, but the print group scored higher on a test of inferences based on the story.

Krendl and Watkins (1983) exposed fifth-grade children to a 15-minute educational television program. They manipulated the purpose of viewing by telling half of the students to watch it for entertainment purposes; the other half were told that it was an educational program and that they should watch it in order to answer questions. Whereas recall of the storyline was the same for both groups (i.e., number of recalled actions, facts, scenes, etc.), the group instructed to view the program for educational purposes responded to the content with a deeper level of understanding; that is, they reported more story and character elements and included more inferential statements about the meaning of the show.

These studies suggest that the perceptions students have about a medium and the purposes they have for viewing influence the amount of effort that they put into the processing of the message and, consequently, the depth of their understanding of the story. The following sections elaborate on the cognitive mechanisms involved in effortful learning with television and examine the interaction of these processes with the characteristics of the medium. Three issues related to the processing of televised information are examined: the relationship between simultaneously pres-

ented auditory and visual information, the processing pace of transient information, and the use of such transient presentations to inform the transformation functions of mental models. For the first of these issues, there is now a considerable amount of cognitive research available; however, there remains little research on the other two issues.

The simultaneous processing of two symbol systems. An important attribute of video is the ability to use both auditory and visual symbol systems. Within the window of cognitive engagement, how do these symbol systems work, independently and together, to influence comprehension and learning with television? Can either symbol system convey the meaning of a presentation? Does the presentation of both at the same time inhibit or facilitate learning?

Baggett (1979) found that either pictorial or linguistic symbol systems alone can carry semantic information, such as a story line. In this study, college students were presented with either a dialogueless movie, *The Red Balloon,* or an experimentally derived, structurally equivalent audio version. They wrote summaries of episodes within the story either immediately after the presentation or after a week delay. An analysis of the summaries by trained raters found that those written immediately after viewing the dialogueless movie were structurally equivalent to those written immediately after listening to the story. Subjects could construct a semantic macrostructure (i.e., summary) from either medium, but information obtained visually was more memorable. Summaries written a week after viewing the movie were judged to be more complete than those written a week after listening to the audio version.

Meaning can be conveyed by either symbol system. However, Baggett (1989) concludes that information presented visually and linguistically is represented differently in memory. She contends that visual representations contain more information and are *bushier.* Whereas the phrase *red leaf* contains only the name of an object and a modifier, a mental representation of a red leaf obtained from a picture carries with it information about size, color, and shape. Also, the visual representation has more *pegs* that can be used to associate it with information already in long-term memory. These additional associations also make it more memorable.

But, it is a significant attribute of video that the auditory and visual symbol systems are presented simultaneously. How does a viewer process information from both of these sources? Two basic hypotheses exist. One possibility is that the simultaneous presentation of audio and visual information competes for limited cognitive resources and that this competition actually reduces comprehension. Another possibility is that information presented with these two symbol systems may work together in some way to increase comprehension.

A number of studies have compared a video program with its decomposed audio and visual presentations to determine the role of these two sources of information, individually and together (Baggett & Ehrenfeucht, 1982,1983; Beagles-Roos & Gat, 1983; Gibbons, Anderson, Smith, Field, & Rischer, 1986; Hayes & Kelly, 1984; Hayes, Kelly, & Mandel, 1986; Meringoff, 1982; Nugent, 1982; Pezdek & Hartman: 1983; Pezdek, Lehrer, & Simon, 1984; Pezdek & Stevens, 1984). In none of these studies did the combination of audio and visual information result in lower recall than recall from either source alone. In most of these studies, the combined use of visual and auditory symbol systems resulted in more recall than visual-only and audio-only presentations. This compels the rejection of the hypothesis that simultaneous presentation of audio and visual information necessarily competes for cognitive resources at the expense of comprehension.

Several of these studies used multiple measures of recall to trace the symbol system source of different kinds of knowledge. In a 1982 study, Meringoff asked 9- and 10-year-old children to draw and talk about their imagery and to make and substantiate inferences about a story, *The Fisherman and His Wife.* Compared to those who heard the story, the children who saw the video drew more details and their pictures were more accurate. Children in the audio groups based their inferences about details on previous knowledge and personal experiences (more like those of children in the control group unexposed to the story), and they were frequently in error relative to the verbal descriptions. Beagles-Roos and Gat (1983) compared animated and audiotape presentations of two stories to groups of first- and fourth-grade children. These researchers found that the explicit story content was learned equally well by both treatment groups. The visual groups recalled more details from the story, did better at a picture sequencing task, and based their inferences on depicted actions. The audio groups more frequently retold the stories using expressive language and based their inferences on verbal sources and prior knowledge.

People can construct a mental representation of the semantic meaning of a story from either audio or visual information alone, but it appears that when presented together each source provides additional, complementary information that retains some of the characteristics of the symbol system of origin. Children recall sounds and expressive language from the audio track and visual details from the visual track. It also appears that the bushier nature of representations derived from the visual symbol systems are better for building mental models of the situation than are representations based on audio-linguistic information. Students listening to an audiotape are more likely to get information for this model from memory. Audio may be sufficient for those knowledgeable of a domain, but visual symbol systems supply important situational information for those less knowledgeable.

These results parallel those for text and pictures. However, the processing of text appears to be driven by the construction of a representation of the linguistic information. Comprehension of video appears to be driven by the processing of visual information. This is apparent from a study by Baggett (1984), who varied the temporal order of audio and visual information within a video presentation on the names and functions of pieces of an assembly kit. In this study, the narration was presented in synchrony or 7, 14, and 21 seconds ahead of or behind the visual presentation. College students performed best on immediate and 7-day delayed tests of recall of the synchronous and 7-second, visual-then-audio presentations. The worst performance was by groups with the audio presented first. This suggests that, in a video presentation, the visual symbol system serves as the primary source of information and that the audio symbol system is used to elaborate it.

The processing of transient information. Another important characteristic of television is that the information it presents can be, and usually is, transient. Comprehension is affected by the pace of this presentation and by its continuity. Wright et al. (1984) used sixteen, 15-minute-long children's television programs that varied in pace and continuity. Pace was defined by these researchers as the rate of scene and character change. Low-continuity programs were those with scenes that were independent and unconnected (i.e., magazine formats). High-continuity programs were those with connected scenes (i.e., stories). These programs were shown to groups of elementary school children whose recall was measured using serration tasks of still pictures from the shows. The children who viewed slow-paced, high-continuity

programs performed better on these tasks. The effect was additive for younger children.

Surprisingly little research has been done on the effect of pace on comprehension, but this is a potentially crucial variable that may distinguish the process of learning with television and other transient media from learning with stable media, such as text. Wright et al. (1984) defined *pace* as a characteristic of the presentation-the amount of information presented per unit of time (i. e., scene and character changes). But from a cognitive perspective, the critical consideration is *cognitive pace-the* amount of information processed per unit of time. From this perspective, the hypothetical unit of information is the chunk-a semi-elastic unit whose size depends on the familiarity and meaningfulness of the information (Miller, 1956; Simon, 1974). A single word may be a chunk in the following list of words: Lincoln, calculus, criminal, address, differential, lawyer, Gettysburg. Rearranged into Lincoln, Gettysburg, address, criminal, lawyer, differential, calculus, the chunk might be larger than one word (e.g., Lincoln's Gettysburg address) but only if the phrase had some meaning in long-term memory. Simon (1974) examines the results of several experiments to conclude that the capacity of short-term memory is five to seven chunks. He also concludes that it takes between 5 to 10 seconds to fixate each chunk in long-term memory. Thus, whereas the amount of time it takes to process information is relatively constant (i.e., one chunk per 5 to 10 seconds), the number of words processed per unit of time depends on the size of the chunk. This, in turn, depends on relevant prior knowledge in long-term memory.

With books, the reader creates chunks of variable word size to effect a reading pace (i.e., words per unit of time) that accommodates the cognitive requirements of comprehension. With television, the pace of presentation (i.e., words or visual elements per unit of time) is not sensitive to the cognitive constraints of the learner; it progresses whether or not comprehension is achieved. The television viewer may be familiar enough with the information to process it at the pace presented, even if it is fast. That is, the viewer's chunks may be large enough so that the cognitive pace of processing words and ideas keeps abreast with the pace at which they are presented. Even if attention waivers and information is missed, knowledge of a familiar domain can be used to fill in the gaps by supplying information from long-term memory. If the viewer has little domain knowledge, the chunk size will be smaller, and the cognitive pace will drop, perhaps below the pace at which ideas are presented. Also, there is less information from long-term memory to compensate for the information that might be missed. Because the information is transient, the viewer can not regress over it to refresh short-term memory. This situation may result in the cascading comprehension failure mentioned by Anderson and Collins (1988). However, for lack of research, these contentions remain speculative, and empirical work in this area is needed.

The discussion above concentrates on the potential problems created by the transient nature of video information. But this transience may have some advantages in the development of dynamic mental models. As mentioned, Greeno (1989) contends that people use mental models to reason through the solution of problems. This is possible because a mental model is considered to be composed of a connected, *runnable* set of objects, or mental entities (Williams, Hollan, & Stevens, 1983). Each of these has an associated representation of its state, a set of parameters, a set of procedures that modify its parameters, and a set of relationships that connect it with

other objects. The model is *run* by means of propagating a change of state in one object over time to the states of connected objects, using the associated procedures and relationships to modify their parameters. Thus the representation is transformed from the current state to some future state. This information is used to make inferences and solve problems (Holland, Holyoak, Nisbett, & Thagard, 1986).

For example, mental models in physics typically include entities that correspond to physical objects that are encountered in the situation, such as blocks, springs, and pulleys (Larkin, 1983). People operate on the mental entities as they would in real time and make inferences about "what would happen to them next" in order to solve physics problems.

Holland et al. (1986) contend that learning a representation of the transition function is the critical goal in the construction of a mental model. The prospect exists that the transient, time-based character of video information could be used to inform the dynamic properties of mental models, such as those in physics. The observation of objects moving along paths, for example, could provide learners with information needed to make estimates of changes in state. This information would not be available with static information, such as that in text. Whereas learners familiar with the domain might be able to supply such dynamic information from memory or use their prior knowledge to infer dynamic properties from static pictures, those novice to a domain may not be able to supply such constructions and might benefit from the dynamic character of televised information. However, as will be discussed in the subsequent section on learning with computers, this information may not be sufficient to overcome

misconceptions that novices frequently bring to tasks, such as those involving the motion of objects (Clement, 1983; di Sessa, 1982; McCloskey, 1983). Again, due to lack of research in this area, these contentions remain speculative.

Summary and implications. This research paints a picture of television viewers who monitor a presentation at a low level of engagement, their moment-to-moment visual attention periodically attracted by salient audio cues and maintained by the meaningfulness of the material. This creates a window of cognitive engagement. Within this window, their processing is sometimes effortless, resulting in the construction of shallow, unelaborate representations of the information presented. However, when viewing with a purpose, people will attend more thoughtfully, constructing more detailed, elaborate representations and drawing more inferences from them.

The visual component of the presentation is particularly memorable, and the representations constructed with it are especially good for carrying information about situations. The auditory symbol systems carry information about sounds and expressive language and help in interpreting the visual information. Auditory symbol systems alone draw primarily on prior knowledge for a construction of the situation model, and this may be problematic for those with little prior knowledge.

Viewers use their prior domain knowledge to process information at the pace presented and supplement information that they may have missed. The transient information in the presentation may be useful in building the dynamic properties of mental models, so that inferences can be made about the phenomena they represent. However, if the topic is unfamiliar, little information exists in long-term memory to supplement viewing. The pace of the presentation may exceed their capacity to process it, and comprehension failure may result.

This knowledge can be used by instructional designers to make media-related decisions. For example, people who are very knowledgeable about a particular

domain can process information at a much faster rate and more strategically with text than they can with audiotape or video, suggesting that text would suffice for these learners. However, people who are novices to a domain are likely to benefit from the ability to slow the rate of information processing, regress over text, and move back and forth between text and pictures as they are presented in books. These same people are more likely to fail at comprehending some portion of a video presentation because their pace of processing information may fall below the pace at which it is presented. Thus, for novices, a more stable medium should be used, or the pace of a video production should be slowed (Kozma, 1986). For learners moderately familiar with a topic, television's symbol systems can supply complementary information, particularly useful in constructing a situation model, and its normal pace will accommodate comprehension. In video productions, the linguistic information should be presented simultaneously with or just following the visual information.

Learning With Computers

So far, media have been described and distinguished from each other by their characteristic symbol systems. Some media are more usefully distinguished by what they can do with information-that is, their capability to process symbols. This is particularly the case for computers, the prototypic information processors. For example, computers can juxtapose, or transform, information in one symbol system to that in another (Dickson, 1985). A learner can type in printed text, and a computer with a voice synthesizer can transform it into speech. The computer can take equations, numerical values, or analog signals and transform them into graphs. Research is reviewed below that shows how the computer can be used to aid students in constructing links between symbolic domains, such as graphs, and the real world phenomena they represent. The research shows that it is the transformation capabilities of the computer, rather than its symbol systems, that are crucial in this regard.

The computer is also capable of *proceduralizing* information. That is, it can operate on symbols according to specified rules, such that a graphic object on the screen can move according to the laws of physics, for example. Research is reviewed below that illustrates the role that this capability can play in aiding learners to elaborate their mental models and correct their misconceptions with the use of microworlds.

Connecting the real world to symbols with MBL. An important part of school learning is acquiring an understanding of the relationship between various symbol systems and the real world they represent. Yet, students are frequently unable to connect their symbolic learning in school to real world situations (Resnick, 1987). The transformational capabilities of the computer can be used to make this connection.

Graphs provide an example of this. Mokros and Tinker (1987) found frequent errors among seventh- and eighth-grade students in the interpretation of graphs. Two patterns were identified. First, there was a strong graph-as-

picture confusion. Half of the students drawing a graph of a bicyclist's speed uphill, downhill, and on level stretches drew graphs representing the hills and valleys rather than speed. In a less striking pattern, 75% of the students responded incorrectly when asked to specify maximum warming or cooling on a graph. About half of the 75 % selected the highest (or lowest) point on the graph as that showing the most rapid change.

Mokros and Tinker (1987) went on to use a microcomputer-based lab (MBL) with 125 seventh and eighth graders for three months. MBL involves the, use of various

p 195

sensors (temperature probes, microphone, motion sensors, etc.) connected to the computer to collect analog data. The computer transforms these data and displays them in real time on the screen as a graph. In a typical unit, the user can turn a heater on for a fixed period, thereby delivering a fixed quantity of thermal energy to a liquid. Using temperature probes interfaced to the computer, the increase and decrease of temperature is instantaneously graphed over time. Mokros and Tinker found a significant increase from pre- to posttests on the interpretation of graphs (from m = 8.3 to m = 10.8 on a 16-item test). Of particular importance was the fact that students made the greatest gains on items sensitive to the graph-as-picture error.

In a similar study, Brasell (1987) used MBL with high school physics students. One group of students spent a class session collecting and observing M13L data in real time (the standard-MBL group). A second group used the MBL equipment to collect data, but it was displayed after a 20-second delay. One control group plotted data with pencil and paper, and another control group engaged in testing only. Brasell found that the posttest scores from the standard-M]3L treatment were significantly higher than scores from all other treatments. The analysis indicated that real-time transformation of data (i.e., the difference between standard-MBL and delayed-MBL) accounted for nearly 90% of the improvement relative to the control. Brasell suggests that unsuccessful students lack appropriate techniques for referring to previous events or experience, and they fail to make explicit links between physical events and the graphed data, even when they are displayed after only a 20-second delay. The transformation capabilities of the computer made the connection between symbols and the real world immediate and direct.

Building mental models with microworlds. Experts in a domain are distinguished from novices, in part, by the nature of their mental models and how they use them to solve problems. The processing capabilities of the computer can help novices build and refine mental models so that they are more like those of experts.

In physics, a series of studies (Chi, Feltovich, & Glaser, 1981; Hegarty, Just, & Morrison, 1988; Larkin, 1983; Larkin, McDermott, Simon, & Simon, 1980) has established that experts have extensive domain knowledge organized into large, meaningful chunks, or schemata, that are structured around the laws of physics. These schemata contain not only information about the laws of physics but also information on how and under which conditions they apply. In other words, they contain both declarative and procedural knowledge.

When encountering a text book physics problem, experts use the objects (e.g., *springs, blocks, pulleys)* and features mentioned in the problem statement to cue the retrieval of one or more relevant schemata (e.g., *force-mass, work-energy).* They construct a mental model, that contains both information that has been explicitly provided by the situation as well as information supplied from memory. These mental models include mental entities that correspond to the physical objects mentioned in the problem *(blocks, pulleys;* Larkin, 1983; Larkin et al., 1980), as well as entities that correspond to the formal constructs of physics that have no direct, concrete referent in the real world *(force, vectors, friction, velocity).* The relationships among these entities correspond to the laws of physics. Experts reason with this model to test the appropriateness of potential quantitative solutions. It is only after this qualitative analysis is complete that the expert will use an equation to derive a quantitative solution to the problem.

Novices represent and use information in this domain in a very different way. Not only do they have less knowledge about physics than do experts, but their knowledge

p 196

is organized quite differently. For some novices, their physics-related knowledge is composed of a set of fragments, or *phenomenological primitives,* that are not connected by formal relationships but are based on real world objects and actions. They evoke these fragments to construct a representation of a particular problem (di Sessa, 1988). Other novices may have coherent and consistent, though erroneous "theories," or misconceptions, of the phenomenon (Clement, 1983; McCloskey, 1983). These may represent procedural relationships that are contrary to established laws of physics, such as: An object remains in motion only as long as it is in contact with a mover, or an object should always move in the direction that it is kicked.

Confronted by a text book problem, novices will use the same surface cues as experts to evoke this information from memory. However, unlike those of experts, the mental models that novices construct with this information are composed primarily of entities that correspond to the familiar, visible objects mentioned in the problem statement (Larkin, 1983). These representations do not contain entities that represent formal physical constructs, such as *force or friction.* Nor do they contain information on physical laws and principles, or this information is inaccurate or incomplete. Thus, the models are insufficient to determine a solution, or the solution that is specified is incorrect.

How do people modify such incomplete and inaccurate mental models to form more accurate, expert-like models? This process is not automatic. Indeed, such misconceptions can be held into adulthood as well as after taking courses in the domain (McCloskey, 1983). Rather, modification of a mental model is triggered by certain conditions, such as the failure of a model to adequately predict or account for phenomena when it is used to achieve some desired goal (Holland et al., 1986). In such cases, a person can drop the current mental model in favor of another, maintain the model but lower confidence in its ability to reliably predict, or modify the model. The latter is most often the goal in school learning. One way a model is modified is by elaborating its situational components. These are the criteria used to evoke and select the appropriate model in response to a particular problem. Another way to modify a model is by changing the transformation rules associated with the situation. Which of these various changes ultimately occurs depends on the accumulated previous success with the model (a model that has been used successfully many times is more likely to be modified than replaced), the perceptual elements of the situation that might allow for differentiation (the existence of salient perceptual elements will be used to refine the selection criteria so that it is used in a somewhat different set of situations), and the future success of alternate models and rules when they compete to explain subsequent situations (modifications in the model that successfully predict subsequent situations are more likely to be retained). Expertise is developed through a series of such differentiations and elaborations as a result of extensive experience within a domain-both successful and unsuccessful.

Now, how might the processing capabilities of computers be used by novices to aid them in building more expert-like models? First, an important attribute of the computer is its ability to symbolically represent entities in ways that might inform mental models. They can graphically represent not only concrete objects but also formal, abstract entities, entities that novices do not normally include in their models. Second, the computer has the important capability of being able to proceduralize the relationships among these symbols. Abstract concepts can be represented in other media, such as text, by symbolic expressions (e.g., f = *ma)* or denoted

in diagrams by arrows, but Greeno (1989) points out that such symbols do not behave like forces and accelerations. With computer models, arrows and other symbols can behave in ways that are like the behavior of forces, velocities, and other abstract concepts. For example, a velocity arrow can become longer or shorter, depending on the direction of acceleration. Furthermore, learners can manipulate these symbols and observe the consequences, successful or otherwise, of their decisions. By implementing their mental models and manipulating these entities governed by the laws of physics, novices may become aware of the inadequacies and inaccuracies of their models. Through a series of such experiences, they can progressively move from initial fragmented, inconsistent, and inaccurate understanding to more elaborate, integrated, and accurate mental models of the phenomena.

This is illustrated in several studies by White (1984, in press), who examined students as they learned principles of Newtonian dynamics within computer-based microworlds. She extended the work of di Sessa (1982), who created a computer-based LOGO environment, called *Dynaturtle,* in which the task was to hit a target through a series of directional "kicks" imparted to the turtle. Di Sessa observed that physics-naive, elementary school students in his study commonly operated with an Aristotelian model of force and motion expressed as: If you impart a force on a moving object, then it will go in the direction last pushed. This Aristotelian notion of force can be contrasted with the Newtonian principle that the motion of an object is the vectorial sum of the forces that have acted on it. An Aristotelian strategy universally used by these students was to wait until the moving turtle was at the same height as the target and give it a 90° kick directly toward the target. The result in this Newtonian environment would be a compromised motion of 45° that would miss the target.

White (1984) analyzed the correct, Newtonian strategy, decomposing it into component principles (i.e., the scalar sum of forces, the vectorial sum of forces, etc.) I and created a series of games that progressively incorporated these component strategies. Each game instantiated both observable objects (e.g., a space ship) and formal physical objects (e-g., a force, represented by a key press). These objects were governed by one of the component Newtonian principles (e.g., combining two forces to increase speed in one direction). The series led up to the target game used by di Sessa. White found that the group of high school physics students who used these games for less than an hour

not only used the Newtonian strategies in the target game but showed significant improvement on transfer verbal force and motion problems. They also performed significantly better on these problems than did a control group of students who attended a physics class but were unexposed to the games.

White and Frederiksen (1990) present a paradigm for the development of a progression of computer models that support conceptual change. The progression leads the learner from simple models to advanced models, increasing in the number of rules, qualifiers, constraints taken into account, and range of problems accommodated. The models allow students to make predictions, explain system function and purpose, solve problems, and receive feedback and explanations. Each is designed to build upon and facilitate transformation from the previous model.

White (in press) applied this progressive paradigm to develop a 2-month curriculum in Newtonian mechanics. This version contained significant improvements in the design. Additional formal constructs from physics were represented by dynamic symbols. For example, a history of the object's speed was represented by a "wake,"

and the vectorial components of forces acting on the object were represented by a "datacross." As the learner applied more force to the object, he or she saw not only the resulting effect on the object as it moved but a dynamic decomposition of the force into its orthogonal vectors (i.e., the datacross) and a dynamic representation of the change in velocity (i.e., its wake). The students were also provided with additional structure, such as a set of possible "laws" to test within the microworld and a set of real word transfer problems. Additional forces, such as friction and gravity, could be introduced into the system. Two classes of sixth graders were assigned to this curriculum for 45 minutes a day, instead of their regular science course. At the end of the period, the groups using the microworld scored significantly better on a range of real world transfer problems than did two classes of sixth graders attending the regular science class. They also scored significantly better on these items than did four classes of high school physics students, including two classes that had just spent 21/2 months studying Newtonian mechanics.

Summary. The studies above examined the processing capabilities of the computer and showed how they can influence the mental representations and cognitive processes of learners. The transformation capabilities of the computer connected the symbolic expressions of graphs to the real world phenomena they represent. Computers also have the capability of creating dynamic, symbolic representations of non-concrete, formal constructs that are frequently missing in the mental models of novices. More importantly, they are able to proceduralize the relationships between these objects. Learners can manipulate these representations within computer microworlds to work out differences between their incomplete, inaccurate mental models and the formal principles represented in the system.

White's research (1984, in press) shows that novice learners within these environments benefit from structured experiences of progressive complexity that help them build and elaborate their mental models. Research by Brasell (1987) and others suggests that such symbolic-operational environments would be particularly powerful if directly connected to real time phenomena. These could help learners connect their more elaborate models to the real world experiences that they can explain.

Learning With Multimedia

This final section is the most speculative. Little research (particularly process research) has been done on learning with multimedia environments, in part because most efforts in the field are focused on development and in part because the field is still evolving. However, multimedia present the prospect that the various advantages of the individual media described above can be brought together in a single instructional environment and strategically used to facilitate learning.

The term multimedia has been around for several decades (Brown, Lewis, & Harclerod, 1973). Until recently, the term has meant the use of several media devices, sometimes in a coordinated fashion (e.g., synchronized slides with audiotape, perhaps supplemented by video). However, advances in technology have combined these media so that information previously delivered by several devices is now integrated into one device. The computer plays a central role in this environment. It coordinates the use of various symbol systems-presenting text and, in another window, presenting visuals. It also processes information it receives, collaborating with the learner to make subsequent selections and decisions.

The following sections review work on two, somewhat different but soon to be integrated, approaches to multimedia environments: *interactive videodisc environ-*

ments and hypermedia environments. The literature reviewed reports on developments within these fields, speculates on the cognitive impact of these environments, and raises issues that must be addressed in future research.

Connecting mental models to the real world with interactive video. Interactive video integrates computer and video technologies in a way that allows both video and computer-generated information to be displayed together. In some implementations, this information is displayed on the same screen and can be overlayed. So, for example, the video could present a view of a boulder rolling down a hill in one window on the screen. The computer could generate force vectors and overlay them on the moving object. In another window, a graph could be generated that plotted velocity or acceleration over time. Alternatively, the student may be given @ workspace within which he or she could compute acceleration or velocity.

The Cognition and Technology Group at Vanderbilt University (1990; Sherwood, Kinzer, Bransford, & Franks, 1987; Sherwood, Kinzer, Hasselbring, & Bransford, 1987) has developed a series of interactive video-based, complex problem spaces (or *macrocontexts)* that are anchored in realistic goals, activities, and situations. These macrocontexts provide semantically rich environments in which students and teachers can collaboratively explore concepts and principles in science, history, mathematics, and literature and use these multiple perspectives to solve realistic problems. The Group contends that the videodisc presentation provides a more veridical representation of events than text and that its dynamic, visual, and spatial characteristics allow students to more easily form rich mental models of the problem situation.

Nationally, a number of interactive videodisc environments are now in the stages of development and formative evaluation. One such environment is *Palenque* (Wilson & Tally, 1990). *Palenque* is intended to be an entertainment and educational exploratory environment for children aged 8-14. With *Palenque,* the viewer becomes a member of an archaeological team of scientists and children exploring ancient Maya ruins in search of the tomb of Pacal, the 12-year-old ruler of Palenque during its heyday.

In an "explore mode," the viewer can use a joystick to engage in "virtual travel;" that is, the video uses a subjective camera perspective to allow the viewer to "see" what he or she would be seeing if he or she were actually there, walking and climbing among the ruins. This is accompanied by a dynamic you-are-here map. The child can use simulated research tools such as a camera, compass, and tape recorder. In the ,'museum mode," the viewer can browse through a database of relevant information including text, still photographs, motion video, graphics, and so on. These are organized into theme "rooms," such as "Maya glyphs" and the "tropical rain forest." In the "game mode," the viewer engages in such activities as putting back together fragmented glyphs and constructing a jungle symphony. Formative evaluation is examining the system's user friendliness, the appeal of the various components, and its comprehensibility.

These systems may be particularly powerful in representing social situations and tasks, such as interpersonal problem solving, foreign language learning, or moral decision making. Situational information needed to understand and solve these semantically rich problems is sometimes difficult to represent by computer alone and can be better represented with video. But, as mentioned earlier (Salomon, 1983), video information alone can easily be processed in a mindless, shallow way, thus

---

reducing the inferences that viewers draw from it. With interactive video, the computer can be used to help the learner analyze the rich information present in a video scene and carefully think through all of the factors that impinge on the problem.

For example, Covey (1990) has created a particularly compelling moral case study, entitled A *Right to Die? The Case of Dax Cowart.* In this environment, students are faced with the real-life dilemma of a young man who, having just returned from the war in Vietnam, is involved in a flaming accident in which he is burned over 60% of his body and loses his sight. In addition, as part of this bum therapy he must be subjected to daily, painful antiseptic washings. He demands to have the treatments discontinued and to be allowed to die. If the treatments are continued, he can be rehabilitated to a functional but disabled life. The student is confronted with an important moral decision: Should the treatments be discontinued?

The goal of the program is not to teach or argue the student toward a specific position but to provide him or her with a moral sensorium within which to explore these issues. Covey (1990) contends that to understand the moral position of another person one must do more than walk in his shoes. One must live in his skin. With this program, which is based on a true case and filmed with the actual people involved, the student can see the patient's treatments and, in effect, "talk" to the patient, the patient's mother, the doctors, a nurse, and a lawyer. The student is guided

through a consideration of the issues of pain and suffering, competence and autonomy, quality of life, and the role of health professionals. Whichever decision the student makes, he or she is presented with contrary information intended to push him or her toward a deeper understanding of his or her position.

Cross-media research on the Dax case study is currently underway to examine the impact of video alone, text alone, and interactive video on the representation and processing of this information and on the moral reasoning of the learners. Also being examined is the interaction between these media and students' prior knowledge, experience, and opinions. Of particular interest will be the social and interpersonal cues embedded in the video information and how these are moderated by computer-generated text and guidance to affect the learners' construction of a model of the situation.

Stevens (1989) shows how these cues can be built into a system and used in problem solving. In this system, a subjective camera view is used to put the learner at the head of a conference table in the role of team leader. The task before the team of programmers is to review and critique program code generated by various members of the team. Critiques can, of course, be done in ways that generate defensiveness and otherwise reduce team productivity, and such incidents are built into the episode as it is played out. The task of the learner/team leader is to manage the meeting and interject comments at appropriate times to facilitate group process. The precise timing and nature of these interjections is left open and up to the learner. Successful behavior within the system must be responsive to social information embedded in the presentation. The learner can interrupt the session at a particular point and use various menus to construct a verbal statement and give it an affective, emotional loading. The feedback is also contextual; an expert system knowledge-base is used to present reactions of the team members as they might be in a real meeting.

Holland et al. (1986) indicate that mental models of social worlds are also filled with misconceptions and stereotypes. Typically, people believe social behavior to be

more predictable at the level of the individual than it is actually. People tend to explain social behavior in terms of dispositions of actors rather than the character of the situation confronting the actor. Interactive video environments, such as the ones above, may help learners build models of social situations and use them to understand social behavior and solve social problems.

Navigating through symbolic expressions with hypermedia. To this point, this article has spent a considerable amount of time discussing the relationship between media and the construction of situation models. Kintsch (1989), however, points out that some texts, such as literary texts, are studied in their own right. In these cases, a major component of the task is to understand a text in the context of other texts and cultural artifacts to which it refers and within which it was constructed. This section describes an implementation of multimedia called *hypermedia* and speculates on its cognitive effects.

Although *hypertext and hypermedia* have become common terms only recently, they are ideas that have been around for several decades. The terms were coined by Nelson (1987/1974) in the sixties, but his thinking was strongly influenced by the earlier work of Bush (1945). As defined by Nelson, *hypertext* is nonlinear text. What it has come to mean in its many emerging implementations is a set of windows on the computer screen that are linked to information in a database (Conklin, 1987). *Hypermedia* is an extension to include a variety of symbolic expressions beyond texts.

These terse definitions can benefit from an illustration. Picture a text document displayed in a window on the computer screen. This document can be searched by various means, including a Boolean key word search using logical functions such as AND and OR. Imagine that the document is an English translation of Plato's *Republic* and that, if desired, the user could display the document in Greek as well as in another window on the screen. In the English version, one could select a word, and the computer could identify its corresponding word in the Greek text; this operation would be reciprocal. There may be other information connected to a word or passage in the text. For example, a passage could be connected to a contemporary scholarly article that comments on it; this article could be retrieved from the database and displayed on the screen. A reference to Homer would allow the user to retrieve and display *The Iliad.* Or, a word could be associated with a dictionary definition, a diagram, a sound, or a bit-mapped, high resolution photograph of an ancient artifact, sculpture, or building. The name of a city or country could be linked to a map of it. The title of a play could be linked to a video enactment of its dramatization that could be displayed in yet another window.

Much of the educational development of hypermedia is occurring in a few universities, such as Project Perseus at Harvard (Crane, 1990), Intermedia at Brown University (Landow, 1989), and Hyperties at the University of

Maryland (Marchionini & Shneiderman, 1988). The domains include the Greek classics, works of English literature, and technical material.

Spiro and Jehng (1990) contend that hypertexts facilitate the application and transfer of complex knowledge to new situations. Such cognitive flexibility requires the representation of knowledge along multiple rather than single conceptual dimensions. The ill-structured nature of complex situations also requires the assembly of representations, rather than the retrieval of an intact schema. According to Spiro and Jehng (1990), hypertext facilitates this cognitive flexibility because it allows a topic to be explored in multiple ways using a number of different concepts or themes. This

p 202

results in the development of integrated, flexible knowledge structures interconnected by criss-crossing conceptual themes that facilitate the use of this knowledge to solve a wide range of problems. Each concept can be subsequently used in many different ways and the same concept can apply to a variety of kinds of situations. The potential cognitive effects of such systems become apparent when one compares their capabilities to the reading behavior of experts as described in the previously mentioned Bazerman (1985) study. These experts read very selectively, making strategic decisions based on a particular purpose and on highly developed schemata f their field. They scanned tables of contents and read parts of articles selectively d in a personally constructed order. At times, they progressed through the text rapidly, and, at other times, they slowed, moving back and forth within and across . This nonlinear reading would certainly appear to be facilitated by the richness information and the nonlinear structure of hypertext.

The process may also be facilitated by an implementation of hypertext that is not widely used, Most current implementations of hypertext systems are search-and-browse systems. The learner is presented with an established database, structured by he author, and is free to navigate through it in whatever way he or she may want. Other systems (e.g., Kozma, in press; Kozma & Van Roekel, 1986; Scardamalia, Bereiter, McLean, Swallow, & Woodruff, 1989) allow learners to add their own information and construct their own relationships, perhaps symbolically representing them by graphic, node-and-link structures. Such systems can be made to correspond to the processes learners use when constructing interrelationships among concepts in real memory, As Salomon (1988) points out, this may prompt learners not only to think about ideas but to think about how they are interrelated and structured. More importantly, they provide an explicit model of information representation that, under certain conditions, learners may come to use as mental models of their thinking.

Beyond the considerable literature that lauds the potential for such systems and describes individual projects, there is little research on hypertext to date. Those studies that have been done (e.g., Gay, Trumbull, & Mazur, in press; Marchionini, 1989; Egan, Remde, Landauer, Lochbaum, & Gomez, 1989) focus on the more rudimentary functions of hypertext (such as search functions) and relatively simple tasks (e.g., identifying specific information in text), rather than learning or problem solving. There are some encouraging preliminary findings in these studies to indicate that hypertext both calls on and develops cognitive skills in addition to those used with standard text, but much more research is needed. The Bazerman (1985) study suggests that much of the reading behavior exhibited by expert physicists is due to their considerable domain knowledge and skill with the medium. Similar research is needed on the impact of domain knowledge and skills in hypertext.

Indeed, in a note of caution, Charney (1987) suggests that some of the very features that make hypertext so appealing may make it more difficult to use for certain students. For example, the nonlinear nature of hypertext requires readers to decide what information to read and in what order; building such sequences is likely to be particularly difficult for readers new to a domain. By comparison, the author-determined sequence of information in text and the use of certain cues to signal structural relationships may particularly facilitate comprehension for novices. Getting lost in hypertext is another potential problem, particularly for novices who lack the extensive schemata that would allow them to easily locate new information within that previously encountered. Finally, lacking domain-based selection criteria, nov-

p 203

ices may end up reading a great deal of material that is not relevant to their purpose. Thus, hypertext seems to hold some promise, but it also poses some challenges, challenges that warrant research in this area.

Summary and implications. Integrated multimedia environments bring together the symbolic and processing capabilities of the various media described above to help learners connect their knowledge to other domains.

Interactive videodisc environments hold the potential for helping learners build and analyze mental models of problem situations, particularly social situations. Hypermedia environments are designed to help the reader build links among texts and other symbolic expressions and construct meaning based on these relationships. Plausible rationales have been given for the expected effectiveness of such environments, but these must be tested, and in some cases serious questions have been raised. Nonetheless, instructional designers will find these to be powerful development environments that have important implications for practice.

For example, these environments may dramatically change the nature of the media decisions made by instructional designers. Until now, the selection of media has been a macrolevel decision. That is, the decision-should video be used or is audiotape sufficient?-has been based on various instructional considerations in balance, and it applies to the entire instructional presentation and to all learners. The desirability of presenting visual information for one component of the task would have to be balanced against the increased cost for the entire presentation.

The structure of these traditional, macrolevel decisions has affected the conduct of media research. The important question for media researchers has been: What is the overall impact of one medium versus another across learners, and is this impact going to be sufficient enough to justify the additional production and delivery costs that might be involved? This is the meta-question that has driven research on media for the past thirty years and has resulted in little understanding of learning with media.

On the other hand, media decisions for integrated multimedia environments will be microlevel decisions. With these environments, it is possible to reconfigure a presentation in response to the needs of a particular learner. The moment to moment selection of appropriate media can respond to specific learner needs and task demands. Audio-linguistic or even text information may be sufficient for most of the presentation or for most learners, but visual information can easily be presented to a particular learner, for a particular segment, at a particular moment, and for a particular purpose.

The macrolevel decision still exists; the cost of such multimedia delivery environments is high, relative to other devices. However, equipment costs are likely to continue to come down, and they are, for the most part, one time costs. Production costs can actually be lower for such systems. Only selected segments need be videotaped; a single segment can be produced based on pedagogical grounds without having to incur the costs of videotaping the entire presentation. Design costs need not go up if the system is used to make these decisions as the interaction progresses so as to avoid the need for programming all possible branches in advance (Stevens, 1989).

A shift from macro- to microlevel design decisions requires an understanding of the moment-by-moment collaboration between a particular learner and the medium. This collaboration raises a different set of questions for the media researcher: What is the prior knowledge of a particular learner? How is this knowledge represented and

structured, and how does the learner operate on it to solve problems? What is the range among learners of such representations and operations? What symbol systems can best represent various components of the task domain? How do these correspond to the way learners represent the task? What skills do the learners have in processing various symbol systems? How do they process various symbol systems together? How can the medium process these in a way that supports the learner?

Many of these questions were addressed in the research reviewed above, and this research can inform microlevel media decisions. However, the fact that these questions are now asked from within an integrated, multimedia environment will raise other, more novel questions-ones not yet addressed in research.

Conclusions

Do media influence learning? Clark (1983) contends that media do not influence teaming under any condition, but the research reviewed in this article suggests that this position must be modified. Some students will learn a particular task regardless of the delivery device. Others will be able to take advantage of a particular medium's characteristics to help construct knowledge.

Various aspects of the learning process are influenced by the cognitively relevant characteristics of media: their technologies, symbol systems, and processing capabilities. For example, the serial processing of linguistic and pictorial information in books is very much influenced by the stability of this technology. Some learners rely on pictures to help construct a textbase and map it onto a model of the situation; others can provide this model from information in memory and do not need pictures or find audio presentations sufficient. The processing of linguistic and visual information in television is very much influenced by the simultaneous presentation of these symbol systems and the information in their codes. Some learners use these to build rich representations of situations,

particularly of their dynamic aspects; others can supply this information from memory, and text or audio presentations suffice. The process of learning with computers is influenced by the ability of the medium to dynamically represent formal constructs and instantiate procedural relationships under the learner's control. These are used by some learners to construct, structure, and modify mental models; other students can rely on prior knowledge and processes, and the use of computers is unnecessary.

However, Clark (1983) contends that, even if there are differences in learning outcomes, they are due to the method used, not the medium. With this distinction, Clark creates an unnecessary schism between medium and method. Medium and method have a more integral relationship; both are part of the design. Within a particular design, the medium enables and constrains the method; the method draws on and instantiates the capabilities of the medium. Some attributions of effect can be made to medium or method, but there is much shared variance between them, and a good design will integrate them. In the various studies cited above, learning was influenced by the methods used, but it was in part because they took advantage of the medium's cognitively relevant capabilities to complement the learner's prior knowledge and cognitive skills. Many of these methods would have been difficult or impossible to implement in other media.

Finally, Clark (1983) calls for a moratorium on media research, but this article provides a rationale for additional research on media. There is a growing understanding of the mechanisms of learning with media, but a number of questions remain, and

the cognitive effects of the more recently developed environments are speculative. Research is needed to extend this understanding.

This research can itself be facilitated by the use of media. Computers provide a unique opportunity to examine learning processes and how these interact with the capabilities of a medium. Particularly useful is the computer's ability to collect moment-by-moment, time-stamped log files of key presses, typed responses, menu selections, and so forth. These data, supplemented by videotapes of students working individually and thinking aloud, can be used to examine the effects of media on learners' mental representations and cognitive processes (Ericsson & Simon, 1984). Videotapes. of several students working together and talking can provide insights into how cognition is shared among students and between students and media (Roschelle & Pea, 1990). The integration of computer and video records will allow for powerful analyses of qualitative data, and the sharing of these analyses among researchers. The examination of the same raw qualitative data by psychologists, anthropologists, and sociologists can bring multiple disciplinary perspectives to bear on media research as well as facilitate the linkage of these knowledge domains that too often go unconnected.

Ultimately, our ability to take advantage of the power of emerging technologies will depend on the creativity of designers, their ability to exploit the capabilities of the media, and our understanding of the relationship between these capabilities and learning. A moratorium on media research would only hurt these prospects.

Notes

[1] Greeno also points out that at least in some cases information in the situation may be used directly without the need to construct and operate on mental representations. Pictures can be considered either as symbolic expressions or as concrete objects in the environment. Pictures as situated objects may be a more efficient source for processing certain kinds of information, quite apart from how that information is represented in memory. See, for example, Larkin and Simon (1987) and Larkin (1989).

[2] It is important to keep in mind that graphic objects, such as those used by di Sessa (1982), may not be symbolic. That is, the objects may not be viewed as having a referent in another domain (e.g., physics), and students may learn to operate on them directly in their own right without taking them to represent concrete objects or physical concepts. The extent to which objects refer to other domains, and thus serve as symbols, should be explicitly addressed in research with symbolic environments.

References

Anderson, D. R., Alwitt, L. F., Lorch, E. P., & Levin, S. R. (1979). Watching children watch television. In G. A. Hale & M. Lewis (Eds.), *Attention and cognitive development (pp.* 331-361). New York: Plenum.

Anderson, D. R., & Collins, P. A. (1988). *The impact on children's education: Television's influence on cognitive development* (Working Paper No. 2). Washington, DC: Office of Educational Research and Improvement.

Anderson, D. R., & Field, D. (1983). Children's attention to television: Implications for production. In M. Meyer (Ed.), *Children and the formal features of Revision* (pp. 56-96). Munich, Germany: Saur.

Anderson, D. R., & Lorch, E. P. (1983). Looking at television: Action or reaction? In J. Bryant & D. R. Anderson (Eds.), *Children's understanding of television* (pp. 1-33). New York: Academic Press.

p 206

Anderson, D. R., Lorch, E. P., Field, D. E., Collins, P., & Nathan, J. (1986). Television viewing at home: Age trends in visual attention and time with television. *Child Development,* 57, 1024-1033,

Anderson, D. R., Lorch, E. P., Field, D. E., & Sanders, J. (1981). The effects of TV program comprehensibility on preschool children's visual attention. *Child Development, 52,* 151-157.

Anderson, R. C., Spiro, R. J., & Anderson, M. C. (1978). Schemata as scaffolding for the representation of information in discourse. *American Educational Research Journal, 15,* 433-440.

Baddeley, A. (I 981). The concept of working memory: A view of its current state and probable future development. *Cognition, 10,* 17-23.

Baggett, P. (1979). Structurally equivalent stories in movie and text and the effect of the medium on recall. *Journal of Verbal Learning and Verbal Behavior, 18,* 333-356.

Baggett, P. (1984). Role of temporal overlap of visual and auditory material in forming dual media associations. *Journal of Educational Psychology, 76,* 408-417.

Baggett, P. (1989). Understanding visual and verbal messages. In H. Mandl &J. Levin(Eds.), *Knowledge acquisition from text and pictures* (pp. 101-124). Amsterdam, The Netherlands: Elsevier

Baggett, P., & Ehrenfeucht, A. (1982). Information in content equivalent movie and text stories. *Discourse Processes, 5,* 73-99.

Baggett, P,, & Ehrenfeucht, A. (1983). Encoding and retaining information in the visuals and verbals of an educational movie. *Educational Communication and Technology Journal, 3](1),* 23-32,

Bayle, E. (1942). The nature and causes of regressive movements in reading. *Journal of Experimental Education, 11,* 16-36.

Bazerman, C. (1985). Physicists reading physics. *Written Communication, 2(l),* 3-23.

Beagles-Roos, J., & Gat, 1. (1983). Specific impact of radio and television on children's story comprehension. *Journal of Educational Psychology, 75(l),* 128-137.

Brandsford, J.D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior,* 11, 717-726.

Brasell, H. (1987). The effect of real-time laboratory graphing on learning graphic representations of distance and velocity. *Journal of Research in Science Teaching, 24(4),* 385-395. Brown, A. L. (1980). Meta-cognitive development and reading. In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 453-481). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18(l),* 32-42,

Brown, J., Lewis, R., & Harclerod, F. (1973). *AV instruction: Technology, media, and method* (4th ed.). New York: McGraw-Hill.

Bush, V. (1945, July). As we may think. *Atlantic Monthly,* pp. 101-108.

Buswell, G. (1937). *How adults read.* Chicago: University of Chicago Press.

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

Clement, J. (1983). A conceptual model discussed by Galileo and used intuitively by physics students. In D. Genter & A. Stevens (Eds.), *Mental models* (pp. 325-340). Hillsdale, NJ: Lawrence Erlbaum Associates.

Charney, D. (1987). Comprehending non-linear text: The role of discourse cues and reading strategies. In J. Smith & F. Halan (Eds.), *Hypertext '87 proceedings* (pp. 109-120). New York: Association for Computing Machinery,

Chi, M. T., Feltovich, P. I,, & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121-152.

Cognition and Technology Group at Vanderbilt University, The. (1990). Anchored instruction and its relationship to situated cognition. *Educational Researcher,* 19(6), 2-10.

p 207

Conklin, J. (1987). Hypertext: An introduction and survey. *IEEE Computer,* 20(9), 17-41. Covey, P. (1990, April). A *right to die?: The case of Dax Cowart.* Paper presented at the annual meeting of the American Educational Research Association, Boston.

Crane, G. (1990). Challenging the individual: The tradition of hypermedia databases. Aca*demic Computing, 6(11),* 36-41.

Dickson, W. P. (1985). Thought-provoking software: Juxtaposing symbol systems. *Educational Researcher, 14(5),* 30-38.

di Sessa, A. (1982). Unlearning Aristotelian physics: A study of knowledge-based learning. *Cognitive Science, 6,* 37-75.

di Sessa, A. (1988). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age* (pp. 49-70). Hillsdale, NJ: Lawrence Erlbaum Associates.

Egan, D. E., Remde, J. R., Landauer, T. K., Lochbaum, C. C., & Gomez, L. M. (1989, April). *Acquiring information in books and SuperBooks.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Ericsson, K., & Simon, H. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Gay, G., Trumbull, D., & Mazur, J. (in press). Designing and testing navigational strategies and guidance tools for a hypermedia program. *Journal of Educational Computing Research.* Gibbons, J ' , Anderson, D. R., Smith, R., Field, D. E., & Rischer, C. (1986). Young children's recall and reconstruction of audio and audiovisual narratives. *Child Development, 57,* 1014-1023.

Goodman, N. *(1976). Languages of art.* Indianapolis, IN: Bobbs-Merrill.

Greeno, J. (1989). Situations, mental models, and generative knowledge. In D. Khahr & K. Kotovsky (Eds.), *Complex information processing* (pp. 285-318). Hillsdale, NJ: Lawrence Eribaum Associates.

Haas, C. (1989). "Seeing it on the screen isn't really seeing it:" Computer writers' reading problems. In G. Hawisher & C. Selfe (Eds.), *Critical perspectives on computers and composition instruction* (pp. 16-29). New York: Teachers College Press.

Halliday, M. A., & Hasan, R. (1976). *Cohesion in English.* London, England: Longman. Hayes, D. S., & Kelly, S. B. (1984). Young children's processing of television: Modality differences in the retention of temporal relations. *Journal of Experimental Child Psychology, 38, 505-514.*

Hayes, D. S., Kelly, S. B., & Mandel M. (1986). Media differences in children's story synopses: Radio and television contrasted. *Journal of Educational Psychology, 78(5),* 341-346.

Hegarty, M., & Just, M. A. (1989). Understanding machines from text and diagrams. In H. Mandl & J. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp. 171-194). Amsterdam, The Netherlands: Elsevier.

Hegarty, M., Just, M. A., & Morrison, 1. R. (1988). Mental models of mechanical systems: Individual differences in qualitative and quantitative reasoning. *Cognitive Psychology, 20,* 191-236.

Holland, J., Holyoak, K., Nisbett, R., & Thagard, P. (1986). *Induction: Processes of inference, learning, and discovery.* Cambridge, MA: MIT Press.

Huston, A., & Wright, J. (1983). Children's processing of television: The informative functions of formal features. In J. Bryant & D. R. Anderson (Eds.), *Children's understanding of television* (pp. 35-68). N.Y: Academic Press

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension.* Newton, MA: Allyn Bacon.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction integration model. *Psychological Review, 95(2),* 163-182.

Kintsch, W. (1989). Learning from text. In L. B. Resnick (Ed.), *Knowing and learning: Essays in honor of Robert Glaser* (pp. 25-46). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kintsch, W., & van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review, 85(5),* 363-394.

Kozma, R. (1986). Implications of instructional psychology for the design of educational television. *Educational Communications and Technology Journal, 34, 11-19.*

Kozma, R. (in press). Constructing knowledge with learning tool. In P. Kommers, D. Jonassen, & T. Mayes (Eds.), *Mindtools: Cognitive technologies for modeling knowledge.* Berlin, Germany: Springer-Verlag.

Kozma, R. B., & Van Roekel, J. (1986). *Learning tool.* Ann Arbor, MI: Arborworks.

Krendl, K. A., & Watkins, B. (1983). Understanding television: An exploratory inquiry into the reconstruction of narrative content. *Educational Communications and Technology Journal, 31(4),* 201-212.

Kuntz, G. C., Drewniak, U., & Schott, F. (1989, April). *On-line and off-line assessment of self regulation in learning from instructional text and picture.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293-323.

Landow, G. P. (1989). Course assignments using hypertext: The example of Intermedia.
Journal of Research on Computing in Education, 21(3), 349-365.

Larkin, J. H. (1983). The role of problem representation in physics. In D. Genter & A. Stevens (Eds.), *Menial models* (pp. 75-98). Hillsdale, NJ: Lawrence Erlbaum Associates.

Larkin, J. H. (1989). Display-based problem solving. In D. Khahr & K. Kotovsky (Eds.), *Complex information processing* (pp. 319-342). Hillsdale, NJ: Lawrence Erlbaum Associates.

Larkin, J. H., McDermott, J., Simon, D., & Simon, H. (1980). Expert and novice performance in solving physics problems. *Science,* 80(4450), 1335-1342.

Larkin, J. H., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11,* 65-100.

Levie, W., & Lentz, R. (1982). Effects of text illustrations: A review of research. *Educational Communication and Technology Journal, 30(4),* 195-232.

Marchionini, G. (1989, April). *Information seeking in electronic encyclopedias.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco. Marchionini, G., & Schneiderman, B. (1988). Finding facts vs. browsing knowledge in hypertext systems. *IEEE Computer, 21(l),* 70-80.

Mayer, R. (1984). Aids to prose comprehension. *Educational Psychologist, 19,* 30-42.

McCloskey, M. (1983). Naive theories of motion. In D. Genter & A. Stevens (Eds.), *Mental models* (pp. 299-324). Hillsdale, NJ: Lawrence Erlbaum Associates.

Meringoff , L. (1982, April). *What pictures can and can't do for children's story understanding.* Paper presented at the annual meeting of the American Educational Research Association, New York.

Meyer, B. J. F. (1975). *The organization of prose and its effects on memory.* Amsterdam, The Netherlands: Elsevier.

Meyer, B. J. F. (1985). Signaling the structure of text. In D. Jonassen (Ed.), *The technology of text* (pp. 64-89). Englewood Cliffs, NJ: Educational Technology Publications. Miller, G. A. (1956). The magical number seven plus or minus two: Some limits of our capacity for processing information. *Psychological Review, 63,* 81-97.

Minsky, M. A. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision* (pp., 211-277). New York: McGraw-Hill.

Mokros, J., & Tinker, R. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching, 24(4),* 369-383.

Nelson, T. (1987/1974). *Computer lib.* Redmond, WA: Tempus Books.

Norman, D. (1989, June). *Cognitive artifacts.* Paper presented at the Workshop on Cognitive Theory and Design in Human-Computer Interaction, Chappaqua, NY.

Nugent, G. C. (1982). Pictures, audio, and print: Symbolic representation and effect on learning. *Educational Communication and Technology Journal, 30(3),* 163-174.

Pea, R. (1990, April). *Distributed intelligence and education.* Paper presented at the annual meeting of the American Educational Research Association, Boston.

p 209

Perkins, D. N. (1985). The fingertip effect: How information-processing technology shapes thinking. *Educational Researcher,* 14(7), 11-17.

Perkins, D. (1990, April). *Person plus: A distributed view of thinking and learning.* Paper presented at the annual meeting of the American Educational Research Association, Boston.

Pezdek, K., & Hartman, E. F. (1983). Children's television viewing: Attention and comprehension of auditory versus visual information. *Child Development, 54,* 1015-1023.

Pezdek, K., Lehrer, A., & Simon, S. (1984). The relationship between reading and cognitive processing of television and radio. *Child Development, 55,* 2072-2082.

Pezdek, K., & Stevens, E. (1984). Children's memory for auditory and visual information on television. *Developmental Psychology, 20,* 212-218.

Pressley, M. (1977). Imagery and children's learning: Putting the pictures in developmental perspective. *Review of Educational Research, 47,* 585-622.

Resnick, L. (1987). Learning in school and out. *Educational Researcher, 16(9),* 13-20.

Roschelle, J., & Pea, R. (1990, April). *Situated learning: Computer-based multimedia analysis of learning activities.* Paper presented at the annual meeting of the American Educational Research Association, Boston.
Rusted, R., & Coltheart, V. (1979). The effect of pictures on the retention of novel words and prose passages. *Journal of Experimental Child Psychology, 28,* 516-524.
Salomon, G. (1974). What is learned and how it is taught' The interaction between media, message, task, and learner. In D. Olson (Ed.), *Media and symbols: The forms of expression, communication, and education* (pp. 383-408). Chicago: University of Chicago.
Salomon, G. (1979). *Interaction of media, cognition, and learning.* San Francisco: Jossey-Bass.
Salomon, G. (1983). The differential investment of mental effort in learning from different sources. *Educational Psychologist, 18(l),* 42-50.
Salomon, G. (1984). Television is "easy" and print is "tough;" The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology, 76(4),* 647-658.
Salomon, G. (1988). Al in reverse: Computer tools that turn cognitive. *Journal of Educational Computing Research,* 4(2), 123-134.
Salomon, G. (1990, April). *If intelligence is distributed, what about the cultivation of individuals' abilities?* Paper presented at the annual meeting of the American Educational Research Association, Boston.
Salomon, G., & Clark, R. (1977). Reexamining the methodology of research on media and technology in education. *Review of Educational Research, 47,* 99-120.
Scardamalia, M., Bereiter, C., McLean, R., Swallow, J., & Woodruff, E. (1989). Computer supported intentional learning environments. *Journal of Educational Computing Research, 5,* 51-68.
Schallert, D. L. (1980). The role of illustrations in reading. In R. Spiro, B. Bruce, & W. Brewer *(Eds.), Theoretical issues in reading comprehension* (pp. 503-524). Hillsdale, NJ: Lawrence Eribaum Associates.
Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Lawrence Erlbaum Associates.
Shebilske, W. L., & Fisher, D. F. (1983). Eye movements and context effects during reading of extended discourse-. In K. Rayner (Ed.), *Eye movements in reading* (pp. 153-179). New York: Academic Press.
Shebilske, W. L., & Reid, L. S. (1979). Reading eye movements, macro-structure and comprehension. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Processing of visible language* (Vol. 1; pp. 97-110). New York: Plenum.
Sherwood, R., Kinzer, C., Bransford, J., & Franks, J. (1987). Some benefits of creating macrocontexts for science instruction: Initial findings. *Journal of Research in Science Teaching,* 24(5), 417-435.
Sherwood, R., Kinzer, C., Hasselbring, T., & Bransford, J. (1987). Macro-contexts for learning: Initial findings and issues. *Applied Cognitive Psychology, 1,* 93-108.

Simon, H. A. (1974). How big is a chunk? *Science, 183,* 482-488.
Snow, R. (1989). Aptitude-treatment interaction as a framework for research on individual differences in learning. In P. Ackerman, R. Sternberg, & R. Glaser (Eds.), *Learning and individual differences (pp.* 13-60). New York: Freeman.
Spiro, R., & Jehng, J. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R. Spiro (Eds.), *Cognition, education, and media* (pp. 163-206). Hillsdale, NJ: Lawrence Ertbaum Associates.
Stevens, S. (1989). Intelligent interactive video simulation of a code inspection. *Communications of the ACM,* 32(7), 832-43.
Stone, D., & Glock, M. (1981). How do young adults read directions with and without pictures? *Journal of Educational Psychology, 73(3),* 419-426.
Vygotsky, L. S. (1978). *Mind in society,* Cambridge, MA: Harvard University Press.
White, B. (1984), Designing computer games to help physics students understand Newton's laws of motion. *Cognition and Instruction, l(l),* 69-108.
White, B. (in press). A microworld-based approach to science education. In E. Scanlon & P. O'Shea (Eds.), *New directions in educational technology.* New York: Springer-Verlag.
White, B., & Frederiksen, J. (1990). Causal model progressions as a foundation for intelligent learning environments, In W. Clancey & E. Soloway (Eds.), *Artificial intelligence and learning environments* (pp. 99-158). Cambridge, MA: MIT Press.

Williams, M. D., Hollan, J. D., & Stevens, A. L. (1983). Human reasoning about a simple physical system. In D. Genter & A. Stevens (Eds.), *Mental models* (pp. 131-154). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wilson, K., & Tally, W. (1990). The "Palenque" Project: Formative evaluation in the design and development of an optical disc prototype. In B. Flagg (Ed.), *Formative evaluation for educational technologies* (pp. 83-98). Hillsdale, NJ: Lawrence Eribaum Associates.

Winn, W. (1989). The design and use of instructional graphics. In H. Mandl & J. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp. 125-144). Amsterdam, The Netherlands: Elsevier.

Wixson, K. (1984). Level of importance of post questions and children's learning from text. *American Educational Research Journal,* 2](2), 419-433,

Wright, J. C., Huston, A. C., Ross, R. P., Calvert, S. L., Tolandelli, D., Weeks, L. A., Raeissi, P., & Potts, R. (1984). Pace and continuity of television programs: Effects on children's attention and comprehension, *Developmental Psychology,* 20(4), 653-666.

Author

ROBERT B. KOZMA is Program Director, National Center for Research to Improve Postsecondary Teaching and Learning and Associate Professor and Associate Research Scientist, School of Education and Center for Research on Learning and Teaching, University of Michigan, Ann Arbor, MI 48109. He specializes in media theory, instructional innovation in higher education, and the design and evaluation of educational software.