Educ Psychol Rev (2006) 18:119–139
DOI 10.1007/s10648-006-9006-3

ORIGINAL PAPER

Promoting Constructive Activities that Support Vicarious Learning During Computer-Based Instruction

Barry Gholson · Scotty D. Craig

Abstract This article explores several ways computer-based instruction can be designed to support constructive activities and promote deep-level comprehension during vicarious learning. Vicarious learning, discussed in the first section, refers to knowledge acquisition under conditions in which the learner is not the addressee and does not physically interact in any way with the source of the content to be mastered. The second section describes cognitive constructivism from the standpoint of schema theory and the work of Bartlett (1932). The next section describes four principles of curriculum design that support constructive processes during vicarious learning and reviews the process of self-explanation and how computer prompted self-explanation supports constructive activities. Research showing the important role that overhearing deep-level reasoning questions plays in supporting constructive activities and deep-level learning is also described. In the next section, vicarious learning supported by deep-level reasoning questions is contrasted with tutoring as one kind of interactive learning. In the final section, some conclusions are drawn, a few empirical issues are discussed, and two caveats are noted.

Keywords Vicarious learning · Constructivism · Self-explanation · Deep-level reasoning questions

Recent advances in computer-based courses (Anderson, Corbett, Koedinger, & Pelletier, 1995; Derry, & Potts, 1998; Holland, Kaplan, & Sams, 1995; Lesgold, Lajoie, Bunzo, & Eggan, 1992) and distance learning (Renwick, 1996; Scardamalia, Bereiter, Brett, Burtis, Calhoun, & Smith-Lea, 1992) have created situations in which learners sometimes find themselves trying to understand course content in settings where they are observers (Brennan &

B. Gholson (✉)
203 Psychology Building,
The University of Memphis, Memphis, TN 38152
e-mail: b.gholson@mail.psyc.memphis.edu

S. D. Craig
University of Pittsburgh, Learning Research and Development Center,
3939 O’hara st, RM 523, Pittsburgh, PA 15260
e-mail: scraig@pitt.edu

Clark, 1996; Cox, McKendree, Tobin, Lee, & Mayes, 1999; Fox Tree, 1999; McKendree, Stenning, Mayes, Lee, & Cox, 1998; Schober & Clark, 1989) rather than active participants. These new educational technologies present numerous challenges for researchers. Consequently, researchers have examined the conditions under which learners can benefit from computer-based instruction when they are only observers e.g., (Lee, Dineen, & McKendree, 1998; McKendree, Stenning, Mayes, Lee, & Cox, 1998). More specifically, the central issue addressed in this article is this: How can computer-based instruction be designed to support knowledge construction processes (Mayer, 1997; Mayer, 2002) when learners cannot physically interact with or control the content they are attempting to master?

Overview of this article

This review examines how vicarious learning might best occur during computer based instruction. The article begins with a brief discussion of vicarious learning during computer-based instruction. The second section begins with a description of schema theory, followed by a brief summary of Bartlett's work on how activated schemas are modified in the presence of new information and inferences. This second section concludes with a discussion of cognitive constructivism. The third section explores cognitive scaffolding for vicarious learning. It contains three subsections. The first subsection explores principles of curriculum design that support deep-level knowledge construction during vicarious learning. The second subsection describes research on self-explanation and how it can support vicarious learning during computer based instruction. The third subsection explores research on how deep-level reasoning questions support vicarious learning. The article's fourth section contrasts vicarious learning using deep-level reasoning questions with interactive learning using an intelligent tutoring system. The final section draws some conclusions, raises a few empirical issues, and notes two caveats.

Vicarious learning during computer-based instruction

Vicarious learning is defined as knowledge acquisition under conditions in which learners are not addressed and are physically passive, in the sense that they do not physically interact (Mayer & Chandler, 2001) in any way with the source of the content they are attempting to master. Vicarious learning's inception in the psychology of human learning dates back to Bandura's (1962) early work on modeling aggression with children. It continued under such labels as observational learning and social learning e.g., (Bandura, 1977, 1986; Lee, Dineen, & McKendree, 1998; Rosenthal & Zimmerman, 1978). Some recent work on vicarious learning dealt with skill learning e.g., (Shebilske, Jordan, Goettl, & Paulus, 1998). Other recent work, however, investigated various manipulations designed to support constructive processes during vicarious learning (Craig, Gholson, & Driscoll, 2002; Craig, Gholson, Ventura, Graesser, & the Tutoring Research Group, 2000; Driscoll, Craig, Gholson, Ventura, Hu, & Graesser, 2003; Fox Tree, 1999; McKendree, Stenning, Mayes, Lee, & Cox, 1998; McNamara, Levinstein & Boonthum, 2004; McNamara & McDaniel, 2004). In this article, we consider only computer-based instruction in which learners have *no control* over the source of the content they are attempting to master. We exclude from consideration even such easily implemented features as learner self-pacing, even though self pacing supports deep-level knowledge construction e.g., (Clark, 1999; Mayer & Chandler, 2001; Mayer, Dow, & Mayer, 2003; Reigeluth, 1999; Sweller, 1999; Van Merrienboer, 1997; Williams, 1996).

Computer-based instruction frequently involves multimedia presentations (Craig, Gholson, & Driscoll, 2002; Mayer, 1997, 1999, 2001; Sweller, 1999) in which learners receive information in more than a single presentation modality (i.e., visual, auditory). Multimedia presentations could be as simple as pictures provided on printed pages along with recorded audio presented via cassette (Mayer, 1997). However, this article focuses on computer-based presentation formats.

Schema theory and cognitive constructivism

Schema theory

This section describes several kinds of schemas proposed in the past 30 years, highlights Bartlett's work on schema reconstruction or modification, and concludes with a review of cognitive constructivism. This section provides the framework for much of the research reviewed in later sections. A schema is an internal representation of the world, an organization of concepts and actions that can be revised by new information about the world (WordNet 2.1. 2005). A variety of different kinds of schemas were described in the past 30 years. According to Sweller and Cooper (1985), even low-level schemas permit people with limited domain knowledge to recognize patterns and respond with appropriate actions. These kinds of schemas represent basic knowledge, but they are structured and complex. Presumably, as domain knowledge increases, higher-level schemas, called object families, are activated. These involve loosely organized collections of knowledge that work together in given situations. Structures within an object family activate each other and behave as a single schema (Chandler & Sweller, 1992; Prawat, 1996; Sweller & Chandler, 1994).

Another type of schema, one that involves a distributed type of memory activation, is called a cognitive field (Derry, 1996). Some patterns of memory activation are made more accessible than others when these schemas are activated (Bransford & Johnson, 1972). Perhaps this is because schemas determine how previously existing knowledge structures are modified by new content. As addressed later (see *Question asking promotes vicarious learning*), the integrative nature of deep-level reasoning questions presumably activates relevant schemas of various sorts, and this activation makes new content easier to process and map onto existing knowledge structures.

Bartlett's work on schema theory

Bartlett studied schemas in the context of geometrical forms and figures (Reynolds & Flagg, 1977) but is best known for his work on story schemas (Bartlett, 1932). He showed that new ideas and inferences that are generated when schemas are activated are often permanently incorporated into schemas. For Bartlett, reconstructing or modifying schemas as better fits to current experience requires only that the relevant schemas be activated in the presence of a new experience that is useful, understandable, and dominant over past experience.

Cognitive constructivism

Constructivist approaches have, over the past decade, shaped the standards for curriculum and instruction in the United States. Examples include the Standards for the English

Language Arts (National Council of Teachers of English NCTE, 1996) and the National Science Education Standards (National Research Council NRC, 1996). A variety of theoretical frameworks based on constructivist epistemology (Kant, 1981, 1990, 1996) have been proposed including Wundt's constructivism (Wundt, 1897; Wundt, 1904), radical constructivism (Graesser, Person, & Magliano, 1995; von Glasersfeld, 1984; von Glasersfeld, 1990), Piaget's constructivist theory (Inhelder & Piaget, 1969; Piaget, 1952, 1962, 1963, 1970a, 1970b, 1973), the Gestalt work on insight learning (Ash, 1985; Bartlett, 1932; Kohler, 1925 work on schema theory, and cognitive constructivism (Jonassen, Peck, & Wilson, 1999; p. 2).

According to cognitive constructivism, teachers and technology play an indirect role in knowledge acquisition. They provide support activities that stimulate learners to think and aquire knowledge, but learners construct knowledge only through what they themselves do (Prawat, 1996). Learners experience new phenomena, interpret experiences in terms of what they already know, reason about new experiences, reflect on the experiences, and reflect on the reasoning process itself. Jonassen et al. (1999), following Bruner (Bruner, 1990), state that these kinds of activities support meaning making. Meaning making is what cognitive constructivism is all about.

Jonassen et al. (1999) argue that, from the constructivist perspective, knowledge results only from active thought processes (Brown, Collins, & Duguid, 1989; Mayer, 1996). Knowledge is anchored in the context in which active thought processes occur. Because meaning exists only in the mind of the knower, and each knower brings a unique set of experiences and perspectives to the learning environment, that environment should offer multiple perspectives on new content. Knowledge results from active thought processes, and meaning making requires that knowledge be applied either verbally, visually, and/or in the auditory channel. Learners need to think about and reflect on the experiences involved in the knowledge construction process.

It should be noted that constructivists do not believe that all meaning is equally valid simply because it is personally constructed. According to Duffy and Cunningham (1996), the litmus test for constructed knowledge is its viability in the knowledge building community. Shared ideas are accepted and agreed upon within that community. That is, meaning is reflected in shared beliefs. If a person's ideas are discrepant from those held by others in the knowledge community, those ideas are rejected as invalid unless new evidence supporting their validity is provided. Constructive processes involve activated schemas operating on data provided by the outside world. Therefore, incorrect meanings and misconceptions can be constructed. Thus, computer-based instruction should include design features that structure the content (Mayer, Mautone, & Prothero, 2002; Sweller, 1999) to cognitively scaffold (Vygotsky, 1978) the learner's new constructions toward the intended (shared) meaning. Ways this cognitive scaffolding can be used to support vicarious learning in computer-based instruction is the focus of the remainder of this article.

Cognitive scaffolding for vicarious learning

This section begins with a brief discussion of instructional design principles (Chandler & Sweller, 1992; Mayer, 1997, 2001; Sweller, 1988, 1999) that support constructive activities. The next subsection reviews research on self-explanation and its implications for vicarious learning (Chi, de Leew, Chiu, & LaVancher, 1994). The final subsection addresses the role of question asking in vicarious learning (Craig, Gholson, Ventura, Graesser, & the Tutoring Research Group, 2000).

Instructional design

Several decades of research led Richard Mayer to propose a cognitive theory of multimedia learning (for reviews see Mayer, 1997, 1999, 2002; Mayer & Chandler, 2001; Sweller, 1988, 1989, Sweller & Chandler, 1994). The cognitive theory of multimedia learning was not developed in the context of vicarious learning. The supporting empirical work was, however, conducted under vicarious learning conditions because there was no physical interaction between the learner and the source of the presentation materials (Mayer, 1997, 2001; Sweller, 1999). At the most basic level, the multimedia learning theory rests on three assumptions: dual channel processing (Baddeley, 1992; Paivio, 1986), limited capacity processing (Mayer & Moreno, 1998; Mousavi, Low, & Sweller, 1995), and active processing (Chandler & Sweller, 1992; Mayer, 2001; Wittrock, 1974; Wittrock, 1990). Dual channel processing means that auditory and visual information are processed using relatively independent channels (Baddeley, 1992; Paivio, 1986). Limited processing capacity means that each channel (auditory, visual) has a limited capacity in working memory that is easily overloaded (Sweller, 1988). Active processing means that learners engage in the active construction of new knowledge (Wittrock, 1974, 1990). The cognitive theory of multimedia learning specifies four instruction design principles that support deep-level knowledge comprehension. The four principles are modality, contiguity (both temporal and spatial), coherence, and redundancy. These principles and their empirical support are well documented e.g., (Mayer, 1997, 1999, 2001; Moreno & Mayer, 2000; Sweller, 1988, 1999), so the description here is brief.

The modality principle says instruction should provide learners' with opportunities to use cognitive processes within both the auditory and visual channels. It is argued that using the two modalities expands the capacity of working memory e.g., (Mayer, 1997, 2001; Penney, 1989; Sweller, 1999). The temporal contiguity principle says that when visual information is presented (e.g., animations), information presented in the auditory channel should be presented simultaneously with information in the visual channel, rather than before or after it (Mayer & Anderson, 1991, 1992). The spatial contiguity principle says that if printed text and diagrams or pictures are used to present information, the text should be embedded in the visual display rather than appear in a different location (Sweller, 1988, 1999). The coherence principle says that when learners are to construct a causal chain (such as learning the steps that occur before lighting strikes), as is required in many scientific systems, causal links should occur in temporal order with no unnecessary information included (Harp & Mayer, 1998; Mayer, 2001; Moreno & Mayer, 2000). The redundancy principle says that information that is redundant interferes with learning. Redundant information increases the demands of working memory when learners attempt to integrate it with other sources of information. This reduces knowledge construction (Chandler & Sweller, 1992; Sweller, 1999; Sweller & Chandler, 1994).

In part, these principles help explain what is called the split attention effect e.g., (Mayer, 1997, 2001; Sweller, 1988, 1999). This effect was described by Sweller (1988) and his colleagues (Chandler & Sweller, 1992; Sweller, 1988; Sweller & Chandler, 1994; Sweller, Chandler, Tierney, & Cooper, 1990). The split attention effect refers to impairments in learning that can result when it is necessary to mentally integrate disparate sources of information before they can be rendered intelligible (Chandler & Sweller, 1992, 1987; Sweller & Cooper, 1985; Ward & Sweller, 1990).

Mayer's theory, then, provides clear principles concerning how to design computer-based instruction to support the meaningful construction of new knowledge (Ertmer & Newby, 1993; Herrington & Standen, 2000; Jonassen, 1991; Kalyuga, Chandler, & Sweller, 1998; Lefrancias, 1997; Mayer, 2001; Wittrock, 1990). Mayer's aim (1997, 2001, 2002), like that of many other specialists e.g., (Derry, 1996; Sweller, 1999), is to replace the behavioral perspective with a more cognitive and constructivist approach. The behavioral perspective, which dominated classroom instruction throughout most of the last century beginning with Thorndike's early work e.g., (Thorndike, 1898, 1903; Thorndike & Woodworth, 1901), views the learner as passively absorbing new knowledge while engaging in rather mindless activities such as drill and practice, rote memorization, and programmed instruction. As indicated above, the cognitive constructivist perspective emphasizes cognitive activities such as reflecting on the materials, searching for alternative perspectives, reflecting on how what is already known relates to the new content, and self-explanation (Chi, 2000; Chi, de Leew, Chiu, & LaVancher, 1994; McNamara, Levinstein, & Boonthum, 2004; McNamara, & McDaniel, 2004; VanLehn, Graesser, Jackson, Jordan, Olney, & Rosé, in press).

Self-explanation promotes vicarious learning

Research indicates that the cognitive activity of self-explanation supports knowledge construction. Chi et al. (1989) asked college students to engage in self-explanation tasks while learning physics concepts and principles. Learners first studied the prose sections of an introductory physics text. They were then given worked-out problems taken from the text. They were asked to explain aloud what they understood from reading each statement in the worked-out problems. The basic finding, which the authors coined the self-explanation effect, was that learners who generated more self-explanations (mean of 15 explanations per problem) while studying the example correctly solved more transfer problems (82% correct) than those who generated few self-explanations (3 self-explanations per problem, 46% correct).

The self-explanation effect was replicated in other domains, including LISP coding (Pirolli & Recker, 1994), electricity and magnetism applied to a mass spectrometer (Ferguson-Hessler & de Jong, 1990), and algebra word problems (Nathan, Mertz, & Ryan, 1994). Related support came from Webb (1989) who reviewed 19 studies involving students learning mathematics and computer science in small groups. Webb's major findings were that (a) student reception of elaborate explanations had little effect on achievement, but (b) student generation of elaborate explanations raised achievement.

The self-explanation effect also proved effective for younger learners. Chi et al. (1994), using a pretest to posttest design, presented eighth graders with a 101-sentence passage on how the circulatory system works. The sentences were presented one at a time, each on a separate page. Those in a self-explanation group were prompted to explain what each sentence meant to them. The prompts, given by the experimenter, were general, simply a reminder to explain what the sentence meant (Chi et al., 1994, p. 451). That is, the prompts were content free, such as "What do you understand from this sentence?, or "What does this sentence mean to you?" (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001, p. 479]. In an unprompted group, learners simply read through the 101-sentence passage two times.

Both the prompted and unprompted groups showed learning gains from pretest to posttest, but the gains were greater for the prompted group (32%) than the unprompted group (22%). The proportional difference in learning gains between the groups was most pronounced on a knowledge application test where the prompted group scores increased 23 percent whereas the unprompted group scores increased just 13 percent.

For the prompted group, Chi et al. (1994) counted as self-explanations *only* those that included inferences that went beyond the content the learner was self-explaining. Inferences made use of content contained in sentences previously provided in the 101-sentence passage or made use of common-sense knowledge. Self-explanations did not include monitoring statements, paraphrases, comprehension, or bridging inferences (for the rationale, (Chi, de Leew, Chiu, & LaVancher, 1994, p. 455). Chi et al. contrasted high self-explainers ($M = 87$ self-explanations) with low self-explainers ($M = 29$). On the application test, the high self-explainers showed greater pretest to posttest gains than the low self-explainers (33% gain vs. 17%). As an explanation for their findings, Chi et al. suggested that linking new content with content from earlier in the text passage was especially helpful. The high self-explainers had a greater proportion of across-sentence links than low self-explainers (1994, p. 475).

More recently, Chi (2000) expanded on the role of inferences generated during the process of self-explaining. Learners may detect a discrepancy between their own mental (model) representation and the model conveyed by the text passage while generating inferences. After a discrepancy is detected, the learner reconstructs the activated schemas to bring them into better correspondence with the text model. An inference that links the current topic to prior knowledge, whether from earlier content in the text or common-sense understanding, involves a reconstruction process. The activated schemas are brought into closer alignment with the new content provided by the text. These processes involve the kinds of constructive activities described by Bartlett (1932) in the context of the activation and reconstruction of schemas (see *Bartlett's work on schema theory*).

Researchers have also examined the value of automatic prompting. Hausmann and Chi (2002) investigated the efficacy of a computer system that prompted college students to self explain. The materials consisted of a subset of 62 sentences taken from the 101-sentence passage used by Chi et al. (1994). Instead of speaking their self-explanations, learners used a keyboard to type their self-explanations, but they understood that the logs would be analyzed later. In Experiment 1, learners were either asked at the outset to generate and type their self-explanations, or they served in a control group where there was no opportunity to type (read-only control). Few self-explanations were generated in the prompted condition (an average of one per learner), and the two groups did not differ on any measure of learning.

In Experiment 2, Hausmann and Chi (2002) compared prompting by a human tutor with automated prompting by computer. The prompts were taken from a set of 69 content-free prompts published by Chi et al. (2001), Appendix E. The human-prompting condition was a Wizard of Oz arrangement (Maulsby, Greenberg, & Mandler, 1993) where the learner and prompter were seated in front of monitors in separate rooms. The learner was not told in advance whether they were interacting with a human or computer. In this condition, the computer automatically presented one sentence at a time from the text on both the learner's and the tutor's monitors. The learner had two options: type a self-explanation or type "ok" to signal they had nothing to say. The learner then hit the submit button, and what was typed appeared on the prompter's monitor. If the prompter then deemed that a prompt was appropriate, one was delivered.

The dialog pattern for each learner in the human prompting condition was yoked to one learner in the automatic prompting condition in terms of when prompts were given. For example, if the human prompter gave a prompt to a particular learner at sentence 15, then the yoked learner also received a prompt at that sentence. Because the human prompter could be sensitive to the learner's knowledge states, prompts could be given when needed. Due to the yoking procedure, which prompts were given, and when they were given, by the computer in the automatic prompting condition was completely arbitrary. Any differences in favor of the human prompting condition over the automatic condition could then be attributed to well timed versus arbitrary prompting (Hausmann & Chi, 2002, p. 10).

The results revealed that both the human and automatic prompting conditions produced significant learning gains (Chi, de Leew, Chiu, & LaVancher, 1994), and there were no differences between them. On average, the learners were prompted 16 times and generated an average of 7 integrative self-explanations. The number of self-explanations learners generated was reliably correlated with learning, but the number of paraphrases was not. This finding was consistent with an earlier claim (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, de Leew, Chiu, & LaVancher, 1994) that paraphrasing is not a constructive activity and does not contribute to overall learning. McNamara and McDaniel (2004) have shown, hovever, that even paraphrase increases learning for those with low domain knowledge. After discussing findings contrasting the number of prompts that were given, by both human tutors (Chi, de Leew, Chiu, & LaVancher, 1994) and computer tutors (Aleven & Koedinger, 2000; Aleven, Koedinger, & Cross, 1999; Renkl, 1997; Wathen, 1997), with the number of self explanations, Hausmann and Chi (2002) concluded that when or where prompts are given is less important than the absolute number of prompts.

Chi et al. (2000) (Chi, 1994), concluded that self-explanations that integrated new to-be-learned content with previously encoded content (i.e., across sentences within the 101-sentence passage) appeared important in promoting deep-level knowledge construction and schema reconstruction. This was because high self-explainers, who learned much more at deeper levels than low self-explainers, had a much higher proportion of self-explanations linking sentences within the text. In fact, it appeared that about 70% of the self-explanations in the Chi et al. (1994) study involved these within-text links, despite the fact that the prompts were content free and not designed to promote such links.

As Chi et al. (1994, p. 475), pointed out, integrating across sentences when generating self-explanations is a strategy that can be taught and is effective e.g., (Bielaczyc, Pirolli, & Brown, 1995; Kant, 1990; King, 1989; King, Staffieri, & Adelgais, 1998; McNamara & McDaniel, 2004; Nathan, Mertz, & Ryan, 1994; Palincsar & Brown, 1984). King et al. (1998), for example, showed that middle-school children could be readily induced to link sentences in the course content and to generate knowledge integration questions during peer tutoring. Hausmann and Chi (2002) have shown that no training is required to elicit self-explanations during computer-based instruction, at least among college students.

Question asking promotes vicarious learning

It has sometimes been proposed that question generation is one of the processing components that underlies comprehension (Collins, Brown, & Larkin, 1980; Graesser, Singer, & Trabasso, 1994; Hilton, 1990; Kintsch, 1998; Olson, Duffy, & Mack, 1985), problem solving (Klahr & Dunbar, 1988; Reisbeck, 1988), reasoning (Williams, 1996), and creativity (Sternberg, 1987). Asking good questions leads to improved comprehension, learning, and memory of the materials (Davey & McBride, 1986; Gavelek & Raphael, 1985; King, 1989, 1990; 1994; King, Staffieri, & Adelgais, 1998; Palincsar & Brown, 1984). It is well documented that students rarely ask questions in the classroom and those they do ask do not involve deep reasoning e.g., (Dillon, 1988; Graesser & Person, 1994; Van der Meij, 1988). Graesser and Person (1994) pointed out that the median number of students' questions per hour in the classroom is 3.0. Because an average classroom contains about 27 students (Dillon, 1988), the number of questions generated by any particular student is low, averaging about .11 (i.e., 3.0/27) question per hour of classroom instruction (Graesser & Person, 1994, p. 105).

The good news is that it is easy to implement strategies that elicit deep-level reasoning questions from college students (Craig et al., 2000) and even from middle-school students (Davey & McBride, 1986; King, 1989; King, 1990, 1994). The training procedures used to induce question generation among middle-school children were usually quite explicit, conducted as part of classroom activity, and required multiple training sessions. Davey and McBride (1986), for example, used five training sessions distributed across two weeks, whereas King (1994) used multiple sessions spread across several weeks.

Inducing question generation from college students was shown to require less training than was used in the studies involving middle school students. Craig et al. (2000) explored whether vicarious-learning procedures could be used to induce question asking (Collins, Brown, & Larkin, 1980; Kintsch, 1998; Palincsar & Brown, 1984) among college students in a brief period of time (about 30 min). Two male animated agents, a virtual tutor and virtual tutee, located on opposite sides of a monitor, discussed a series of eight computer-literacy topics. For each topic, a relevant picture was also located on the monitor between the virtual tutor and virtual tutee. As shown below, the procedures were successful.

During acquisition, learners either overheard the virtual tutee carry on a dialog with the virtual tutor, or they overheard a more monologue-like discourse. In both conditions, each topic began with a brief information delivery by the virtual tutor. Then, in the monologue-like condition, the virtual tutee asked one broad question that provided a context for what followed, and the virtual tutor answered in a monologue-like discourse that presented all the information on that topic. In the dialog condition, each brief information delivery was followed by a lively series of conversational exchanges. The virtual tutee asked a series of deep-level reasoning questions (Bloom, 1956; Graesser & Person, 1994), a total of 66 across the eight topics. The virtual tutor answered each question. The deep-level reasoning question stems were taken from six categories in a question taxonomy presented by Graesser and Person (1994, pp. 110–111): comparison, interpretation, causal antecedent, causal consequent, instrumentale/procedural, and enablement. The exact words, phrases, and sentences spoken by the virtual tutor in response to the virtual tutee's questions were identical in the dialog and monologue-like conditions for each topic. Immediately following acquisition, free-recall questions on the discourse content of two of the topics were administered.

Next participants were administered a transfer task in which they were presented with a series of eight new computer literacy topics and the opportunity to ask questions on each. Participants were told that (a) the tutor would briefly deliver information on each topic, (b) they (the learners) could direct queries to the experimenter that would help them understand the topic, and (c) the experimenter would answer each question. All queries were audio recorded and transcribed for analysis. Only the computer-controlled virtual tutor and a picture relevant to the particular topic were on the monitor during the transfer task. After the first brief information delivery, the learners continued their queries until they said they were finished. This was followed by a brief information delivery on the next topic, etc., until all eight topics had been presented. Free-recall questions on two of the transfer topics were then administered.

In the transfer task, learners in the dialog condition took more conversational turns (dialog $M = 30$, monologue-like $M = 19$), and generated more queries (dialog $M = 37$, monologue-like $M = 27$). Additionally, those in the dialog condition generated about twice as many questions that involved deep-level reasoning than those in the monologue-like condition. Conversely, those in the monologue-like condition generated a greater proportion of shallow-level reasoning questions than those in the dialog condition (Craig et al., 2000, p. 248). In answering the two free-recall questions following transfer, students in the dialog condition also wrote more propositions ($M = 23$) than those in the monologue-like condition ($M = 17$).

Given the differences in the number of conversational turns, number of questions, and kinds of questions that were asked, it was not surprising that students in the dialog group learned more in the transfer task and outperformed those in the monologue-like group. An unexpected finding was that, following acquisition, learners in the dialog condition also outperformed those in the monologue-like condition on the free-recall questions (dialog $M = 24$ propositions, monologue-like $M = 20$ propositions). The difference was unexpected, and only marginally significant, but the effect size (Cohen's d) was 0.44. Thus, with respect to understanding the conditions that facilitate vicarious learning, an issue that remained to be determined was do students learn (i.e., recall) more when they simply overhear question-rich dialog discourse or when they overhear the same content presented in monologue-like discourse?

At the time there was at least some indirect evidence that students do learn more when overhearing dialog than when overhearing monologue. For example, Fox, Tree (1999) reported that over-hearers (vicarious learners) performed better on a referential communication task while overhearing dialogs than while overhearing monologues. Moreover, (McKendree, Good, & Lee, 2001) reported that over hearers who listened to a question-rich dialog outperformed those who overheard a dialog with few questions.

Driscoll et al. (2003) followed up on the Craig et al. (2000) study which had shown a trend in the free-recall data favoring the dialog condition. The question posed by Driscoll and colleagues was whether the effect would prove more robust if more precise measures were used. In the Craig et al. study the between-subject variability was extreme, ranging from learners who wrote only one or two brief sentences on each of the two free-recall question to others who wrote more than a standard page on each question. Thus, in order to obtain more precise measures of the learners' performances, discourse type (dialog vs. monologue-like) was a within-subject variable in the Driscoll et al. (2003) research. Two male animated agents, a virtual tutor and virtual tutee, engaged in dialog and monologue-like discourse. Each learner overheard four computer literacy topics discussed in dialog format and four in monologue-like format. In the monologue-like format, the virtual tutee asked one question about each of the four topics, but in the dialog format he asked a total of 33 deep-level reasoning questions. As in the Craig et al. (2000) study, the exact same words, phrases, and sentences were spoken by the virtual tutor in both conditions. Following this exposure, each participant was given two free-recall test questions, one probing a topic overheard in dialog format and the other probing a topic overheard in the monologue-like format.

In the Craig et al. (2000) study, the data were scored only for the total number of propositions written on each free-recall question, with no further evaluation of the informational content. Thus, in order to add further precision, the free-recall propositions obtained by Driscoll et al. (2003) were classified into three categories: relevant, related, and irrelevant. Propositions written by learners on a given topic that matched (or paraphrased) those spoken by the virtual tutor on that specific topic were classified as *relevant*. Any propositions that matched (or paraphrased) the content spoken by the virtual tutor on any of the other seven topics were classified as *related* to the content of the topic being scored. Any other propositions written by the learner were classified as *irrelevant* (e.g., meta cognitive comments such as "I don't know much about computers," or propositions concerning the tutorial content that were false). Analyses of the data (Driscoll et al., 2003, Experiment 1) revealed that learners wrote more relevant propositions on free-recall questions probing topics overheard in dialog format ($M = 15$ propositions) than they wrote on questions probing topics overheard in monologue-like format ($M = 9$ propositions). There were no differences in the number of related or irrelevant propositions, with learners in each condition averaging about 2.5 of each kind.

Driscoll et al. (2003) designed Experiment 2 to explore several features of dialog that could be responsible for the effects obtained in Experiment 1. It was possible that the virtual tutee's questions in the dialog condition improved vicarious learning because they (a) provided concept repetition (Fox Tree, 1999), (b) furnished signaling devices similar to headings in printed text e.g., (Hartley & Trueman, 1985; Loman & Mayer, 1983; Lorch & Lorch, 1995), (c) were questions *per se* (McKendree, Good, & Lee, 2001), or (d) were deep-level reasoning questions that activated relevant schemas and provided a coherent context (Bloom, 1956; Duffy, Shinjo, & Myers, 1990; Gernsbacher, 1997; Graesser, Millis, & Zwaan, 1997) for the content spoken by the virtual tutor on his next turn. Thus, the study included four dialog conditions as a between-subject variable, with discourse type (dialog vs. monologue-like) as a within-subject variable in each.

Deep-level reasoning questions (Bloom, 1956; Graesser & Person, 1994) were used in one condition, whereas a second condition involved only shallow-level reasoning questions that mostly required one-word answers (Graesser & Person, 1994, pp. 110–111). An example of a deep-level reasoning question is "How does the CPU use RAM when you are running an application?." An example of a shallow-level reasoning question is "Does the CPU use RAM when you are running an application?". In a simple assertion condition, the shallow-level reasoning questions were transformed into simple assertions spoken by the virtual tutee. For example, "The CPU uses RAM when you are running an application." In a "monologue" condition, the virtual tutee asked only one question per topic (as in the monologue-like condition), and the simple assertions spoken by the virtual tutee were spoken by the virtual tutor. Except in this monologue condition, in which the simple assertions were spoken by the virtual tutor, the words, phrases, and sentences spoken by the virtual tutor were exactly the same in both the dialog and monologue-like conditions.

The rationale for the four conditions was the following. If concept repetition enhances vicarious learning, then performance in all four dialog conditions should exceed that exhibited in the monologue-like conditions. If the tutee's contributions to the dialog functioned as signals, similar to headings in printed text, then deep-level reasoning questions, shallow-level reasoning questions, and simple-assertions should all produce differences when dialog conditions were compared to the monologue-like conditions. If questions *per se* facilitate vicarious learning from overheard dialog, then both the deep-level reasoning questions and shallow-level reasoning questions should produce differences in favor of the dialog conditions, whereas the simple assertions and monologue conditions should not. Finally, if deep-level reasoning questions were the key feature of the overheard dialog (Driscoll et al., 2003, Experiment 1), then only learners overhearing the deep-level reasoning questions were expected to show enhanced learning when the dialog conditions were compared to the monologue-like conditions.

Analyses contrasting the number of propositions written on dialog vs. monologue-like discourse in each dialog condition yielded a significant difference only in the deep-level reasoning questions condition. The mean number of relevant propositions that learners wrote on the topic overheard in dialog format was 17. But, for the topic overheard in monologue-like format, they wrote a mean of 11 propositions. Discourse type did not differ in any of the remaining dialog conditions.

The virtual tutee's deep-level reasoning questions apparently provided a context for the incoming content by activating relevant schemas and making it easy to integrate new content into them. When the learners detected discrepancies between those schemas and the new content (Bartlett, 1932; Chi, 2000), they reconstructed the schemas to bring them into correspondence with that content. Although the learners had low domain knowledge at the outset of the experimental sessions, even low-level schemas permit learners to recognize patterns and respond appropriately. During the course of the session, as the amount of domain knowledge increased, higher-level schemas were activated.

Vicarious learning vs. interactive learning

This section begins with a brief review of the efficacy of human tutoring. This is followed by a subsection on an intelligent tutoring system called Auto Tutor, a subsection that contrasts vicarious learning while observing tutoring sessions with interactive learning, and a final subsection that contrasts vicarious learning using deep-level reasoning questions with interactive learning.

One question that received little previous attention concerns how much knowledge is acquired by vicarious learners when compared to active participants in computer-based instruction formats designed to promote constructive learning processes (Mayer, 2001; Sweller, 1999; Wittrock, 1990). Craig, Driscall, and Gholson (2004) explored this question in the context of an intelligent tutoring system called Auto Tutor (Graesser, Chipman, Haynes, & Olney, in press; Graesser, Person, Harter, & the Tutoring Research Group, 2001; Graesser, Wiemer-Hastings, Wiemer-Hastings, Harter, Person, & the Tutoring Research Group, 2000). There is evidence that the gains from tutoring, when compared to classroom instruction, are generally in the range of 0.4 to 2.0 standard deviation units, depending on the expertise of the tutor (Bloom, 1956, 1984; Cohen, Kulik, & Kulik, 1982; Corbett, 2001; Graesser & Person, 1994).

Human tutoring

A meta-analysis was performed (Cohen, Kulik, & Kulik, 1982) on a large sample of studies that compared human tutoring to standard classroom instruction. Most of the tutors in the studies were untrained and had only moderate domain knowledge. They were peer tutors, cross-age tutors, or paraprofessionals, not accomplished professionals. The average learning gain was 0.4 standard deviation units when compared to various control conditions such as re-reading text or standard classroom activities. The 0.4 gain translates into about half a letter grade. Bloom (1984) reviewed evidence showing that accomplished human tutors when compared to classroom controls produce gains of about 2.0 standard deviation units, or about two letter grades.

AutoTutor

AutoTutor implements the tutoring strategies of paraprofessionals (Graesser & Person, 1994). These strategies mostly involve trying to get the learner to fill in important information, with AutoTutor attempting to correct any misconceptions that are detected (Graesser, Person, & Magliano, 1995). AutoTutor's approach to tutoring was inspired by the constructivist approach to learning described in an earlier section (Chi et al., 1994; Jonassen, Peck, & Wilson, 1999; Piaget, 1952; 1970a; VanLehn, Jones, & Chi, 1992). Various versions of AutoTutor tutor college students on computer literacy and on Newtonian physics. Learning gains obtained in tutoring sessions with versions of AutoTutor generally range from about 0.6 to 1.5 standard deviation units when compared to various controls. In fact, one version of AutoTutor that tutors college students on Newtonian physics produced pretest to posttest learning gains comparable to those produced by physics professors who were experienced tutors (VanLehn et al., in press, Experiments 1 and 5).

AutoTutor, which serves as a conversational partner with the learner, was constructed by the Tutoring Research Group (TRG) at the University of Memphis. An agent on the monitor displays facial expressions and some gesturing while conversing with the learner. AutoTutor begins each topic with an *information delivery*, followed by a *question* for the learner to answer. For each topic, the TRG constructed an *ideal answer* to each question. The ideal answer was then decomposed into sentences containing key concepts called *expectations*. Using latent semantic analysis e.g., (Graesser et al., 2001; Graesser et al., 2000; Landauer & Dumais, 1997), AutoTutor assesses the learners' progress by comparing learners' contributions to the content of each expectation. It builds upon the learners' contributions to the dialog by assuring that each expectation is covered for each topic. Once a given expectation is covered, AutoTutor moves on to another one until all the expectations are covered for that particular topic. This is followed by a brief *summary* before AutoTutor moves on to the next topic.

Vicarious learning while observing interactive tutoring sessions

As part of two larger experiments, Craig et al. (2004), Experiments 1 and 2 contrasted pretest to posttest gains of learners who interacted directly with AutoTutor on 12 computer literacy topics with yoked controls in vicarious learning conditions. Two 24-item multiple-choice tests were used to assess learning gains. The two tests were counterbalanced as pretests and posttests. The visual and auditory contributions of AutoTutor and the contributions of the learner were recorded and presented to one (yoked) learner in each of the vicarious conditions. Whereas learners assigned to both the interactive groups and the vicarious groups showed learning gains from pretest to posttest, those in the interactive groups outperformed those in the vicarious groups in both Experiments 1 and 2.

Vicarious learning with deep-level reasoning questions

Craig, Sullins, Witherspoon, and Gholson (in press), Experiments 1 and 2 contrasted pretest to posttest learning gains across five instructional conditions in Experiment 1 and two (relevant) conditions in Experiment 2. In Experiment 1, there was one interactive condition and four vicarious concitions. In one condition, called "interactive," learners interacted directly with AutoTutor on 12 computer literacy topics. One vicarious condition involved the recorded video and audio that was obtained from each session in the interactive condition. Each recorded session was presented to a yoked control. This condition was called "yoked-vicarious."

The second vicarious condition involved presenting a monologue containing the content of the ideal answer plus the expectations (recall, the expectations were sentences containing key concepts from each ideal answer) to the learner on each topic. These sentences were presented using the same voice engine and on-screen agent that were used in the interactive and yoked-vicarious conditions. The number of sentences in the ideal answers ranged from five to eight across the 12 computer literacy topics, with the same number of expectations included for each topic. Thus, learners were given a monologue presentation of between ten and 16 sentences on each topic. This condition was called "monologue-vicarious." Each topic began with the same information delivery and concluded with the brief summary used in the interactive and yoked vicarious conditions.

In the third vicarious condition, each sentence in the ideal answer was preceded by a deep-level reasoning question. These questions were designed to activate relevant schemas and provide a coherent context for the content of the sentence that followed. The questions were asked by a second, distinct voice engine, but only one agent, the same one used in the other conditions, was located on the monitor. After presentation of the sentences in the ideal answer, each preceded by a question, the expectation sentences were presented, but the expectation sentences were not preceded by questions. That is, the expectation sentences were

Table 1 Group means and standard deviations for pretest and posttest scores, along with group pretest to posttest effect size (Cohen's *d*)

Group	Pretest		Posttest		Cohen's *d*
	M	*SD*	*M*	*SD*	
Interactive	5.65	1.87	10.62	3.63	1.72
Yoked vicarious	6.46	2.43	9.83	3.73	1.07
Ideal-answer vicarious	5.96	2.16	11.11	3.52	1.76
Half-questions vicarious	6.37	1.80	10.89	3.52	1.62
Full-questions vicarious	6.70	1.87	13.30	4.30	1.99

presented in monologue format. This was called the "half-questions vicarious" condition. In the final vicarious condition, each sentence in both the ideal answer and the expectations was preceded by a deep-level reasoning question. This was called the "full-questions vicarious" condition. The information delivery and brief summaries presented in the interactive and other vicarious conditions were presented in the half-questions and full-questions vicarious conditions.

The means and standard deviations of each condition's pretest and posttest scores, along with each condition's pretest to posttest effect size (Cohen's *d*) are presented in Table 1. Results revealed that learners in the full-questions vicarious group outperformed those in each of the other four groups. The mean pretest to posttest change score for the full-questions group was 7, compared to 5 in the interactive group and 3 in the yoked vicarious group. The scores for the interactive group and the yoked vicarious group were similar to those obtained previously in similar conditions (Craig, Gholson, & Sullins, 2004).

Although we were unable to locate any findings in the literature related to the results of Craig et al. (in press). Experiment 1, the findings were partially replicated in Craig et al. (in press). Experiment 2, also in the domain of computer literacy. In That study, learners in a deep-level reasoning questions group outperformed those in an interactive group. Those in the former group showed average pretest to posttest learning gains of 7 and an effect size of 2.29, whereas those in the interactive group showed pretest to posttest gains of only 4 with an effect size of 1.39.

As part of a larger study in the domain of Newtonian physics, the generality of the deep-level reasoning questions effect is presently being investigated (Witherspoon, Sullins, Craig, Brittingham, Lam, & Gholson, in preparation). Two conditions from this study are relevant to the present discussion. In the interactive group, college students interacted with Auto Tutor throughout the learning session. In the deep-level reasoning questions group, Auto Tutor only presented students with deep-level questions and ideal answers. Learning gains were evaluated using a pretest and posttest taken from (VanLehn, Graesser, Jackson, Jordan, Olney, & Rosé, in press) using questions from the Force Concept Inventory (Hestenes, Wells, & Swackhammer, 1992) designed to address learners' understanding of Newton's first, second, and third laws. Although the study is not yet complete, both groups are showing comparable learning gains form pretest to posttest. To put the preliminary findings in a larger context, recall that Auto Tutor delivers learning gains on the Force Concept Inventory (and other measures) comparable to those produced by physics professors who were experienced tutors (VanLehn et al., in press, Experiments 1 and 5).

Summary and conclusions

There are a number of easily implemented computer-based design features that support knowledge construction processes during vicarious learning by scaffolding the learner's constructive activities. Curriculum designers preparing course content for computer-based instruction should implement four features specified by Mayer's (Chandler & Sweller, 1992; Mayer, 1997; Mayer & Chandler, 2001; Sweller, 1999; Sweller & Chandler, 1994) theory that are relevant to their curriculum design whenever possible. Those principles were described earlier in the section called *Instructional design*.

Requesting vicarious learners to self-explain using content-free prompts promotes constructive activities that support deep-level knowledge construction (Chi, 2000; Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, de Leew, Chiu, & LaVancher, 1994; Hausmann & Chi, 2002). The caveat here is that in the Chi et al. (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, de Leew, Chiu, & LaVancher, 1994) research, a human did the prompting, whereas in the Hausmann and Chi experiment (Hausmann & Chi, 2002), Experiment 2, the learners did not know whether prompts were delivered by a human or a computer. In addition, learners in the Chi et al. (1994) and the Housmann and Chi (2002), studies knew their self-explanations were being recorded for later analysis. How effective prompts would be in getting learners to generate genuine self-explanations if they knew the prompts were generated by a computer and their own self-explanations were not being recorded remains to be determined. Prompts that support text links across sentences (Chi 1994) would probably also support constructive activities. It should be pointed out, however, that we failed to locate any related research that used prompts designed to support text links cf. (Davey & McBride, 1986; King, 1989, 1990, 1994; Schuh, 2003). Whether they work as well, or perhaps better, than content free prompts remains to be determined. Research contrasting the use of content free prompts with those designed to support text links across sentences has not been reported. In any case, designing computer-based instruction that includes prompting is, as Hausmann and Chi (2002) demonstrated, easy to implement.

Modeling question asking early in the presentation of course content improves the quality of questions generated later by learners e.g., (King, 1990, 1994; Palincsar & Brown, 1984). The questions induce learners to generate many deep-level reasoning questions in transfer tasks (Craig et al. 2000) and learn more than those in control conditions e.g., (Craig et al., 2000; Davey & McBride, 1986; Driscoll et al., 2003; King et al., 1998). Similarly, preceding each sentence (or main idea) of the course content with deep-level reasoning questions clearly supports constructive processes and vicarious learning. Presumably, these questions activate relevant schemas, support discrepancy detection, and promote schema reconstruction processes (Bartlett, 1932; Chi, 2000). Curriculum designers with domain expertise should not find it difficult to lace deep-level reasoning questions e.g., (Bloom, 1956; Craig et al., 2004; Graesser & Olde, 2003; Graesser & Person, 1994; Otero & Graesser, 2001) into computer-based course content.

The deep-level reasoning questions effect clearly holds in the domain of computer literacy and probably in the domain of Newtonian physics. Unfortunately, we failed to locate any published research contrasting any kind of vicarious learning with interactive learning. There is a caveat here. Both the deep-level reasoning questions and the course content that followed each question were spoken by voice engines in all of the research to date. Whether the same learning gains would be achieved if deep-level reasoning questions were presented as on-screen text before course content was printed is still uninvestigated. Answers to these and related questions await future research.

Acknowledgements The research for this review article was supported by a grant from the Institute of Educational Sciences at the Department of Education (R305H050169). This review was also partially supported by grants from the National Science Foundation (SBR 9720314, REC 0106965, REC 0126265, ITR 0325428), and the DoD Multidisciplinary University Research Initiative (MURI) administered by ONR under grant

N00014–00–1–0600 awarded to the Tutoring Research Group (visit http://www.autotutor.org). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Department of Education, ONR or NSF.

References

Aleven, V., & Koedinger, K. R. (2000). The need for tutoring dialog to support self-explanation. In C. P. Rose, & F. Freedman (Eds.), *Building dialog systems for tutorial applications. Papers from the 2000 AAAI Fall Symposium* (pp. 65–73). Menlo Park, CA: AAAI Press.

Aleven, V., Koedinger, K. R., & Cross, K. (1999). Tutoring answer explanation fosters learning with understanding. In S. P. Lajoie, & M. Vivet (Eds.), *Artificial Intelligence in Education, Open Learning Environments: New Computational Technologies to Support Learning, Exploration, and Collaboration, proceedings of AIED-99* (pp. 199–206). Amsterdam: IOS Press.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences*, *4*, 167–207.

Ash, M. G. (1985). Gestalt psychology: Origins in Germany and reception in the United States. In C. Buxton (Ed.), *Points of View in the Modern History of Psychology*. New York: Academic Press.

Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559.

Bandura, A. (1962). Social learning through imitation. In M. R. Jones (Ed.), *Nebraska Symposium of Motivation* (pp. 211–269). Lincoln: University of Nebraska Press.

Bandura, A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bartlett, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge, England: Cambridge University Press.

Bielaczyc, K., Pirolli, P. L., & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction*, *13*, 221–252.

Bloom, B. S. (1956). *Taxonomy of education objectives: The classification of educational goals. Handbook I: Cognitive domain*. New York: McKay.

Bloom, B. S. (1984). The 2-sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, *13*, 4–16.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1482–1493.

Bransford, 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.

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

Bruner, J. (1990). *Acts of Meaning*. Cambridge, MA: Harvard University press.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*, 233–246.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In Glaser, R. (Ed.), *Advances in Instructional Psychology* (pp. 161–238). Hillsdale, NJ: Erlbaum.

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

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

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25*, 471–533.

Clark, R. C. (1999). *Developing technical training* (2nd edn.). Washington, DC: International Society for Performance Improvement.

Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, *19*, 237–248.

Collins, A., Brown, J. S., & Larkin, K. M. (1980). Inference in text understanding. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension* (pp. 385–407). Hillsdale, NJ: Erlbaum.

Cooper, G., & Sweller, J. (1987). The effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, *79*, 347–362.

Corbett, A. T. (2001). Cognitive computer tutors: Solving the two-sigma problem. *User Modeling: Proceedings of the Eighth International Conference* (pp. 137–147).

Cox, R., McKendree, J., Tobin, R., Lee, J., & Mayes, T. (1999). Vicarious learning from dialogue and discourse. *Instructional Science*, *27*, 431–458.

Craig, S. D., Driscoll, D., & Gholson, B. (2004). Constructing knowledge from dialog in an intelligent tutoring system: Interactive learning, vicarious learning, and pedagogical agents. *Journal of Educational Multimedia and Hypermedia*, *13*, 163–183.

Craig, S. D., Gholson, B., & Driscoll, D. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology*, *94*, 428–434.

Craig, S. D., Gholson, B., & Sullins, J. (2004). Should we question them?: An investigation into the role of deep questions in vicarious learning environments. In J. Nall, & R. Robson (Eds.), *Proceedings of E-learn 2004: Wold Conference on E-learning in Corporate, Government, Healthcare, & Higher Education* (pp. 1836–1840). Norfolk, VA: Association for the Advancement of Computing in Education.

Craig, S. D., Gholson, B., Ventura, M., Graesser, A. C., & the Tutoring Research Group. (2000). Overhearing dialogues and monologues in virtual tutoring sessions: Effects on questioning and vicarious learning. *International Journal of Artificial Intelligence in Education*, *11*, 242–253.

Craig, S. D., Sullins, J., Witherespoon, A., & Gholson, B. (in press). The deep-level reasoning questions effect: The role of dialog and deep-level reasoning questions during vicarious learning. *Cognition & Instruction*.

Davey, B., & McBride, S. (1986). Effects of question generating training on reading comprehension. *Journal of Educational Psychology*, *78*, 256–262.

Derry, S. J. (1996). Cognitive schema theory in the Constructivist debate. *Educational Psychologist*, *31*, 163–174.

Derry, S. J., & Potts, M. K. (1998). How tutors model students: A study of personal constructs in adaptive tutoring. *American Educational Research Journal*, *35*, 65–99.

Dillon, J. T. (1988). *Questioning and Teaching: A Manual of Practice*. New York: Teachers College Press.

Driscoll, D., Craig, S. D., Gholson, B., Ventura, M., Hu, X., & Graesser, A. (2003). Vicarious learning: Effects of overhearing dialogue and monologue-like discourse in a virtual tutoring session. *Journal of Educational Computing Research*, *29*, 431–450.

Duffy, T. M., & Cunningham, D. J. (1996). Constructivism: Implications for the design and delivery of instruction. In D. H. Jonassen (Ed.), *Handbook of Research for Educational Communications and Technology* (pp. 170–198). New York: Simon Schuster Macmillan.

Duffy, S. A., Shinjo, M., & Myers, J. L. (1990). The effect of encoding task on memory for sentence pairs varying in causal relatedness. *Journal of Memory and Language*, *29*, 27–42.

Ertmer, P. A., & Newby, T. J. (1993). Behaviorism, cognitivism, & Constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, *6*, 50–72.

Ferguson-Hessler, M. G. M., & de Jong, T. (1990). Studying physics texts: Differences in study processes between good and poor solvers. *Cognition and Instruction*, *7*, 41–54.

Fox Tree, J. E. (1999). Listening in on monologues and dialogues. *Discourse Processes*, *27*, 35–53.

Gavelek, J. R., & Raphael, T. E. (1985). Metacognition, instruction, and the role of questioning activities. In D. L. Forrest-Presslet, G. E. MacKinnin, & T. G. Waller (Eds.), *Metacognition, Cognitive, and Human Performance*: Vol. 2 (pp. 103–136). Orlando, FL: Academic Press.

Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes*, *23*, 265–304.

Graesser, A. C., Baggett, W., & Williams, K. (1996). Question-driven explanatory reasoning. *Applied Cognitive Psychology*, *10*, S17–S32.

Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (in press). Auto Tutor: An intelligent tutoring system with mixed initative dialog. *IEEE Transactions in Education*.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. In J. T. Spence, J. M. Darley, & D. J. Foss (Eds.), *Annual Review of Psychology* (pp. 163–189). Palo Alto, CA: Annual Reviews.

Graesser, A. C., & Olde, B. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology*, *95*, 524–536.

Graesser, A. C., & Person, N. (1994). Question asking during tutoring. *American Educational Research Journal*, *31*, 104–137.

Graesser, A. C., Person, N., Harter, D., & the Tutoring Research Group (2001). Teaching tactics and dialog in Auto Tutor. *International Journal of Artificial Intelligence in Education*, *12*, 257–279.

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-on-one tutoring. *Applied Cognitive Psychology*, *9*, 495–522.

Graesser, A., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *3*, 371–395.

Graesser, A. C., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., Person, N., & the Tutoring Research Group. (2000). Using latent semantic analysis to evaluate the contributions of students in Auto Tutor. *Interactive Learning Environments*, *8*, 129–148.

Harp, S., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*, 414–434.

Hartley, J., & Trueman, M. (1985). A research strategy for text designers: The role of headings. *Instructional Science*, *14*, 99–155.

Hausmann, R. G. M., & Chi, M. T. H. (2002). Can a computer interface support self-explaining? *Cognitive Technology*, *7*, 4–14.

Herrington, J., & Standen, P. (2000). Moving from an instructivist to a constructivist multimedia learning environment. *Journal of Educational Multimedia and Hypermedia*, *9*, 195–205.

Hestenes, G., Wells, M., & Swackhammer, G. (1992). Force Concept Inventory. *The Physics Teacher*, *30*, 141–158.

Hilton, D. J. (1990). Conversational processes and causal explanation. *Psychological Bulletin*, *107*, 65–81.

Holland, V. M., Kaplan, J. D., & Sams, M. R. (Eds.). (1995). *Intelligent Language Tutors*. Hillsdale, NJ: Erlbaum.

Inhelder, B., & Piaget, J. (1969). *The Early Growth of Logic in the Child*. New York: Norton.

Jonassen, D. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm? *Educational Technology Research and Development*, *39*, 5–14.

Jonassen, D., Peck, K., & Wilson, B. (1999). *Learning with Technology: A Constructivist Perspective*. Upper Saddle River, NJ: Prentice Hall, Inc.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors*, *40*, 1–17.

Kant, I. (1981). *Grounding for the Metaphysics of Morals*. In J. W. Ellington (Ed. and Trans.), Indianapolis, IN: Hackett Publishing Company. (Original work published 1785).

Kant, I. (1990). Critique of pure reason. In J. M. D. Meiklejon (Ed. and Trans.). Buffalo, NY: Prometheus Books. (Original work published 1781).

Kant, I. (1996). Critique of practical reason. In T. K. Abbott (Ed. and Trans.). Amherst, NY: Prometheus Books. (Original work published 1788).

King, A. (1989). Effects of self-questioning training on college students' comprehension of lectures. *Contemporary Educational Psychology*, *14*, 366–381.

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal*, *27*, 664–687.

King, A. (1994). Guiding knowledge construction in the classroom: Effect of teaching children how to question and explain. *American Educational Research Journal*, *31*, 338–368.

King, A., Staffieri, A., & Adelgais, A. (1998). Mutual peer tutoring: Effects of structuring tutorial interaction to scaffold peer learning. *Journal of Educational Psychology*, *90*, 134–152.

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. New York: Cambridge University Press.

Klahr, D., & Dunbar, K. (1988). Duel search space during scientific reasoning. *Cognitive Science*, *12*, 1–48.

Kohler, W. (1925). *The Mentality of Apes*. London: Routledge & Kegan Paul.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Lee, J., Dineen, F., & McKendree, J. (1998). Supporting student discussions: It isn't just talk. *Education and Information Technologies*, *3*, 17–29.

Lefrancias, G. R. (1997). *Psychology for Teachers* (9th edn.). Belmont CA: Wadsworth.

Lesgold, A., Lajoie, S., Bunzo, M., & Eggan, G. (1992). SHERLOCK: A coached practice environment for an electronics troubleshooting job. In J. H. Larkin, & R. W. Chabay (Eds.), *Computer-Assisted Instruction and Intelligent Tutoring Systems: Shared Goals and Complementary Approaches* (pp. 201–238). Hillsdale, NJ: Erlbaum.

Loman, N. L., & Mayer, R. E. (1983). Signaling techniques that increase the understandability of expository prose. *Journal of Educational Psychology*, *75*, 402–412.

Lorch, R. F., Jr., & Lorch, E. P. (1995). Effects of organizational signals on text processing strategies. *Journal of Educational Psychology*, *87*, 537–544.

Maulsby, D., Greenberg, S., & Mandler, R. (1993). Prototyping an intelligent agent through Wizard of Oz. In S. Ashlund, A. Henderson, E. Hollnagel, K. Mullet, & T. White, (Eds.) *Proceedings of the INTERCHI'93 Conference on Human Factors in Computing Systems* (pp. 277–284). Amsterdam: IOS Press.

Mayer, R. E. (1996). Learners as information processors: Legacies and limitations of educational psychology's second metaphor. *Educational Psychologist*, *31*, 151–162.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, *32*, 1–19.

Mayer, R. E. (1999). Multimedia aids to problem-solving transfer. *International Journal of Educational Research*, *31*, 611–623.

Mayer, R. E. (2001). *Multimedia Learning*. New York: Cambridge University Press.

Mayer, R. E. (2002). Using illustrations to promote constructivist learning from science text. In J. Otero, J. A. Leon, & A. C. Graesser (Eds.), *The Psychology of Science Text Comprehension* (pp. 333–356). Mahwah, NJ: Lawrence Erlbaum Associates.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, *83*, 484–490.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, *84*, 444–452.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology*, *93*, 390–397.

Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology*, *95*, 806–812.

Mayer, R. E., Mautone, P., & Prothero, W. (2002). Pictorial aids for learning by doing in a multimedia geology simulation game. *Journal of Educational Psychology*, *94*, 171–185.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*, 312–320.

McKendree, J., Good, J., & Lee, J. (2001). *Effects of Dialogue Characteristics on Performance of Overhearers*. Presented at International Conference on Communication, Problem-solving and Leaning, June 2001, Strathclyde.

McKendree, J., Stenning, K., Mayes, T., Lee, J., & Cox, R. (1998). Why observing a dialogue may benefit learning. *Journal of Computer Assisted Learning*, *14*, 110–119.

McNamara, D. S., Levinstein, I. B., & Boonthum, C. (2004). iStart: Interactive strategy training for active reading and thinking. *Behavior Research Methods, Instruments, & Computers*, *36*, 222–233.

McNamara, D. S., & McDaniel, M. (2004). Suppressing irrelevant information: Knowledge activation or inhibition? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 465–482.

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*, 358–368.

Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology*, *92*, 117–125.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, *87*, 319–334.

Nathan, M. J., Mertz, K., & Ryan, R. (1994). Learning through self-explanation of mathematics examples: Effects of cognitive load (poster). Presentation to the American Educational Research Association (AERA) annual meeting.

National Council of Teachers of English (NCTE) (1996). *Standards for the English language arts*. Urbana, IL: National Council of Teachers of English.

National Research Council (NRC) (1996). *National Cience Educational Standards*. Washington, D.C.: National Academy Press.

Olson, G. M., Duffy, S. A., & Mack, R. L. (1985). Question asking as a component of text comprehension. In A. C. Graesser & J. B. Black (Eds.), *The Psychology of Questions* (pp. 219–228). Hillsdale, NJ: Erlbaum.

Otero, J., & Graesser, A. C. (2001). PREG: Elements ofa model of question asking. *Cognition & Instruction*, *19*, 143–175.

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. Oxford, England: Oxford University Press.

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

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition*, *17*, 398–422.

Piaget, J. (1952). *The Origins of Intelligence*. Madison, CT: International University Press.

Piaget, J. (1962). *Play, Dreams and Imitation in Childhood*. New York: Norton.

Piaget, J. (1963). *The Origins of Intelligence in Children*. New York: Norton.

Piaget, J. (1970a). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's Manual of Child Psychology* (pp. 703–732). New York: John Wiley & Sons.

Piaget, J. (1970b). *Genetic Epistemology*. New York: Columbia University Press.

Piaget, J. (1973). *To Understand is to Invent*. New York: Grossman.

Pirolli, P., & Recker, M. (1994). Learning strategies and transfer in the domain of programming. *Cognition and Instruction*, *12*, 235–275.

Prawat, R. S. (1996). Constructivism, modern and postmodern. *Educational Psychologist*, *31*, 215–226.

Reigeluth, C. M. (1999). *Instructional Design Theories and Models*. Mahwah, NJ: Erlbaum.

Reisbeck, C. K. (1988). Are questions just function calls? *Questing Exchange*, *2*, 17–24.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*, *21*, 1–29.

Renwick, W. (1996). *Informative Technology and the Future of Post Secondary Education*. Paris: OECD.

Reynolds, A. G., & Flagg, P. W. (1977). *Cognitive Psychology*. Cambridge, MA: Winthrop Publishers, INC.

Rosenthal, R. L., & Zimmerman, B. J. (1978). *Social Learning and Cognition*. New York: Academic Press.

Scardamalia, M., Bereiter, C., Brett, C., Burtis, P. J., Calhoun, C., & Smith-Lea, N. (1992). Educational applications of a network communal database. *Interactive Learning Environments*, *2*, 47–71.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, *21*, 211–232.

Schuh, K. L. (2003). Knowledge construction in the learner-centered classroom. *Journal of Educational Psychology*, *95*, 426–442.

Shebilske, W., Jordan, J., Goettl, B., & Paulus, L. (1998). Observation versus hands-on practice of complex skills in dyadic, triadic, and tetradic training-teams. *Human Factors*, *40*, 525–540.

Sternberg, R. J. (1987). Questioninig and intelligence. *Questing Exchange*, *1*, 11–13.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285.

Sweller, J. (1999). *Instructional Design in Technical Areas*. Australian Council for Educational Research: Melbourne.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, *4*, 295–312.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. *Journal of Experimental Psychology: General*, *119*, 176–192.

Sweller, J., & Cooper, M. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*, 59–89.

Tarmizi, R., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology*, *80*, 424–436.

TechSmith Camtasia (Version 3.0) (Computer software). (2001). East Lansing, Michigan: TechSmith Corporation.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review*, Monograph Suppl., 2(8).

Thorndike, E. L. (1903). *Educational Psychology*. New York: Macmillan.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of the other. *Psychological Review*, *8*, 247–261, 381–395, 556–564.

Van der Meij, H. (1988). Constraints on question asking in classrooms. *Journal of Educational Psychology*, *80*, 401–405.

Van Merrienboer, J. J. G. (1997). *Training Complex Cognitive Skills*. New Jersey: Educational Technology Publications.

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rosé, C. P. (in press). Natural language tutoring: A comparison of human tutors, computer tutors and text. *Cognitive Science*.

VanLehn, K., Jones, R. M., & Chi, M. T. C. (1992). A model of the self-explanation effect. *The Journal of the Learning Sciences*, *2*, 1–59.

von Glasersfeld, E. (1984). An introduction to radical constructivism. In P. Watzlawick (Ed.), *The invented reality*. New York: Norton.

von Glasersfeld, E. (1990). An exposition of constructivism: Why some like it radical. In R. B. Davis, C. A. Maher, & N. Noddings (Eds.), Constructivist views on teaching and learning of mathematics. *Journal for Research in Mathematics Education: Monograph 4*, 19–30.

von Glasersfeld, E. (1995). *Radical Constructivism: A Way of Knowing and Learning*. London: Falmer Press.

Vygotsky, L. S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.

Ward, M., & Sweller, J. (1990). Structuring effective worked out examples. *Cognition and Instruction*, *7*, 1–39.

Wathen, S. H. (1997). *Collaborative Versus Individualistic Learning and the Role of Explanation*. Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh, PA.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Education Research*, *13*, 21–39.

Williams, M. D. (1996). Learner control and instructional technologies. In D. H. Janassen (Ed.), *Handbook of Research for Educational Communications and Technology* (pp. 957–983). New York: Macmillan.

Witherspoon, A., Sullins, J., Craig, S. D., Brittingham, J., Lam, C., & Gholson, B. (in preparation). *The Role of Deep-Level Reasoning Questions in the Learning of Newtonian Physics*.

Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist*, *1*, 87–95.

Wittrock, M. C. (1990). Generative processes of comprehension. *Educational Psychologist*, *24*, 345–376.

WordNet 2.1. (2005). Princeton University.

Wundt, W. (1897). *Outlines of Psychology*. In C. H. Judd, (Ed. and Trans.). Leipzig, Germany: Voss.

Wundt, W. (1904). *Principles of Physiological Psychology*. In E. Titchner (Ed., Trans.). London: Swan Sonnenschein. (Original work published 1874).