

## ABSTRACT

Title of Document: PEER COLLABORATION: THE ROLE OF  
QUESTIONS AND REGULATORY PROCESSES  
IN CONCEPTUAL-KNOWLEDGE LEARNING

A. Fielding Ince Winters, PhD 2009

Directed By: Professor Patricia Alexander  
Department of Human Development, University of  
Maryland

Peer collaboration and questioning are two pedagogical methods currently used under the assumption that they facilitate conceptual understanding in science classrooms. However, the literature on peer collaboration reveals many contextual factors that influence the success of peer learning, particularly for ill-structured tasks, and little research has been conducted on whether or how questions help students learn about complex science topics. This study investigated the impact of peer collaboration and reasoning questions on high-school students' (N = 133) conceptual-knowledge learning, through analysis of their regulatory learning processes as they studied the circulatory system using a hypermedia encyclopedia. Outcome variables were a measure of students' conceptual knowledge learning (pretest to posttest) and peers' collaborative discourse, which was collected via audiotape during the learning session. Data analysis consisted of quantitative analyses of variance of students'

conceptual knowledge learning in peer and questioning conditions, and qualitative analysis of students' collaborative regulatory discourse. Results revealed variable approaches to collaboration and the task and variable success at conceptual-knowledge learning across pairs. Successful peer learners employed a variety of regulatory behaviors such as taking notes and summarizing to a greater degree than unsuccessful collaborating students, who tended to spend a large proportion of their time off-task. Students who answered an inferential reasoning question spent much of their time looking for a verbatim answer from the environment, often to the detriment of their learning. The results of this study reveal a number of factors that may be related to the success of collaboration and question-answering, including an accurate perception of the task goal; enough relevant prior knowledge about the topic to use a non-linear hypermedia environment effectively; and enough time to collaborate and learn. This study contributes to the literature on collaboration and question-answering by demonstrating the potential pitfalls of these methods and elucidating potential targets for support to bolster the efficacy of these methods.

PEER COLLABORATION: THE ROLE OF QUESTIONS AND REGULATORY  
PROCESSES IN CONCEPTUAL-KNOWLEDGE LEARNING

By

A. Fielding Ince Winters

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2009

Advisory Committee:

Professor Patricia A. Alexander, Chair

Dr. Ann Battle, Assistant Director, Institute for Child Study

Professor William Holliday

Professor Robert Mislevy

Professor Allan Wigfield

## Table of Contents

CHAPTER I: INTRODUCTION.....	1
Pedagogical Practices for Conceptual-Knowledge Learning.....	3
Peer Learning .....	3
Questioning .....	4
Theoretical Framework .....	8
Theoretical Perspectives on Peer Learning .....	8
Social Cognitive Theory.....	10
Social-Cognitive Theory and Peer Learning.....	11
Other Models and Perspectives .....	14
The Pintrich (2000) Framework.....	14
Research on Collaborative Self-Regulated Learning.....	15
Contextual Framework.....	18
Complex Systems .....	19
Computer-Based Learning Environments, Multimedia, and Hypermedia.....	20
Present Study.....	21
Research Questions and Hypotheses .....	23
Operational Definitions.....	26
CHAPTER II: LITERATURE REVIEW .....	28
Literature Review: Peer Collaboration .....	29
What Do We Know About The Effect of Peer Collaboration on Learning? ....	29
What Do We Know About The Role of Peer Collaboration for Fostering Science Learning? .....	36
Studies with Well-Structured Tasks.....	39
Studies with Ill-Structured Tasks.....	44
Peer Collaboration in science: What do we need to know? .....	55
Literature Review: Questioning .....	56
What Do We Know About The Role of Questioning in Learning? .....	56
Deep-Reasoning Questions .....	61
What Do We Know About Questioning for Fostering Science-Learning?.....	64
Summary .....	69
Questioning in Science: What do we Need to Know? .....	69
Learning with Hypermedia CBLEs.....	71
Theories of Learning with Multimedia and Hypermedia.....	72
Multimodal Theories.....	73
Instructional Theories .....	73
Learning with Hypermedia .....	78
Hypermedia Learning and the Current Study .....	80
CHAPTER 3: METHOD .....	84
Participants.....	84
Research Design.....	85
Materials.....	85

Conceptual-Knowledge Measure .....	85
Hypermedia Learning Environment.....	89
Deep-Reasoning Questions .....	90
Procedure .....	91
Coding and Scoring.....	97
Conceptual-Knowledge Measure .....	97
Peers' Regulatory Discourse .....	99
Question Responses.....	103
Interrater Agreement .....	104
CHAPTER 4: RESULTS.....	105
Descriptive Statistics.....	105
Research Question Results.....	107
Research Question 1 .....	107
Research Question 2.....	109
Research Question 3.....	126
Research Question 4.....	127
Research Question 5.....	128
CHAPTER 5: DISCUSSION.....	131
Peer Collaboration.....	133
Reasoning Questions.....	140
Collaborating on Question Answering.....	143
Conclusion .....	146
Implications and Contributions.....	148
Limitations .....	153
Future Research.....	155
REFERENCES .....	171
APPENDICES .....	160
Appendix A: Conceptual-Knowledge Pretest and Posttest.....	160
Appendix B: Reasoning Question.....	163
Appendix C: Scoring Rubric for Pretest and Posttest Conceptual Knowledge ...	164
Appendix D: Coding Scheme for Peer Discourse.....	166
Appendix E: Scoring Rubric for Answers to Deep-Reasoning Questions.....	170

## List of Tables

Table 1	Matrix of Instructions, by Condition.....	93
Table 2	Experimental Procedure for Questioning and No Questioning Conditions, for Peer and Individual Learners.....	95
Table 3	Summary of Coding and Scoring Procedures.....	98
Table 4	Means, Standard Deviations, Skewness, and Kurtosis Values of Study Measures.....	106
Table 5	Analysis of Covariance for Peer and Question Conditions.....	109
Table 6	Mean, Standard Deviation, and Median Percentages of Collaborative Self-Regulatory Learning Processes Overall Audiotaped Pairs and by School Type.....	110
Table 7	Correlations between Percentage of Collaborative Regulatory Processes and Gain and Posttest Scores.....	112
Table 8	Raw Frequency and Percentages of Coded Utterances for Larger and Smaller Gain.....	114
Table 9	Sample Size, Mean Ranks, and Mann-Whitney U Value for Proportions of Utterances Related to Forethought, Performance, Monitoring, and Motivation.....	130

## List of Figures

Figure 1	Example of coded transcript from Winters and Azevedo (2005).....	102
Figure 2	Mean posttest scores by experimental condition and school type.....	108
Figure 3	Number of total and audiotaped collaborative pair gain combination types.....	113
Figure 4	Interaction between school type and peer learning condition on mean question response score.....	128

## CHAPTER 1: INTRODUCTION

Science education has received increased scrutiny in the last decade as data from international studies, such as the *Programme for International Student Assessment* (PISA) show American students lagging behind other industrialized countries in science achievement (Lemke et al., 2001; Lemke et al., 2004). The most recent data from the 2003 PISA showed that U.S. high-school students scored below the international average for science literacy (Lemke et al., 2004). Further, the latest *Nation's Report Card on Science* (2005) reports that nearly half of high-school seniors (46%) were below even the basic level of achievement for scientific knowledge about critical concepts and skills (Grigg, Lauko, & Brockway, 2006). These recent data reflect a trend that has been evident for the past decade or more (e.g., National Center for Education Statistics [NCES], 1998). In an effort to address such concerns in the last decade, the American Association for the Advancement of Science (AAAS; 1993, 1997) and National Research Council (NRC; 1996) developed science-education standards based on research on effective science pedagogy and content.

According to the AAAS (1993) and the NRC (1996) standards, students should have a solid foundation of concepts about physical and biological science by the beginning of high school. One of the focal points of these science-education standards is fostering students' learning and understanding of conceptual-knowledge as opposed to their acquisition of mere factual information. Conceptual knowledge includes the interaction of a person's declarative and procedural knowledge, and it demonstrates his or her understanding or "knowing why" about a particular topic (Byrnes, 2001; Chi & Ohlsson, 2005; Ryle, 1949). Declarative knowledge is composed of facts, and it is often

defined as “knowing that” about some object or idea, whereas procedural knowledge is composed of skills related to “knowing how” to do something (Ryle, 1949). Procedural knowledge in the domain of science involves the processes of science inquiry (NRC, 1996); skills related to scientific thinking and reasoning (Dunbar & Fugelsang, 2005); skills of argumentation and justification (Driver, Newton, & Osborne, 2000); and many domain-specific process skills, such as classification, measurement, tool selection, and graphing (AAAS, 1993; NRC, 1996).

As such, there has been increased interest in methods of learning that support conceptual understanding. These pedagogical methods for fostering conceptual-knowledge learning have been included in educational standards (see the Teaching Standards, NRC, 1996). Yet, some of these methods warrant further research to determine not only if, but how they affect conceptual-knowledge learning.

This dissertation study investigated two pedagogical practices intended to promote conceptual-knowledge learning: *peer learning* and *questioning*. While these pedagogical approaches have been supported by decades of research (e.g., Thorndike, 1938; Winne, 1979) the results are far from conclusive about mechanisms and efficacy of peer learning (see O’Donnell, 2006; Webb & Palinscar, 1996), and questioning (e.g., Pressley et al., 1992) in the classroom. The results from this study help elucidate how these methods may promote conceptual-knowledge learning by investigating not only the product of learning, but also the regulatory processes students engage in while they are learning collaboratively, either with a reasoning question to guide them, or without. It also theoretically and methodologically extends the research on peer collaboration and

questioning to include analysis of students' collaborative regulatory learning processes and their connection to students' conceptual-knowledge learning.

### Pedagogical Practices for Conceptual-Knowledge Learning

#### *Peer Learning*

Peer learning has been touted in the standards (Teaching Standards E & D, NRC, 1996), as well as by educational researchers for the past decade for its positive role in classroom-based learning (e.g., Johnson & Johnson, 1991; Rosenshine & Meister, 1994; Slavin, 1996). Peer learning refers to a variety of teaching strategies that involve students working together in pairs or small groups to accomplish a mutual educational goal or task (O'Donnell, 2006; Webb & Palinscar, 1996). However, while a well-researched topic in the education field, *peer learning* is not a term that is consistently well-defined within the research literature, as can be seen by the variety of other terms it encompasses, including collaborative learning, cooperative learning (Slavin, 1995), and peer tutoring (Graesser & Person, 1994).

Broadly defined, a peer-learning situation is one in which “two or more people learn, or attempt to learn something, together” (Dillenbourg, 1999, p.2). As Dillenbourg (1999) notes, there are many different ways to operationally define *more people*, *learn something*, and *together*, which results in a wide variety of potential peer learning situations. For example, under this definition, peer learning could include a larger group of people learning interactively through computer-mediated means over the course of one year (Dillenbourg, 1999), while at the same time describing a pair of students working together on a mutual task face-to-face for one or two class periods (e.g., Winters & Azevedo, 2005). The latter situation can be more specifically described by the term peer

collaboration, or collaborative learning. Peer collaboration is the term used to refer to peer-learning situations in which students work together, face-to-face, in a classroom context toward a shared understanding called convergence (Webb & Palinscar, 1996). Cooperative learning is a particular type of peer collaboration in which the collaborative context is structured under specific models for the students, such as *student-teams-achievement-division* or *Jigsaw* (Slavin, 1980). Peer tutoring is not considered a type of peer collaboration, as it occurs in a very particular circumstance, when a more able peer is partnered with a less able peer to promote learning (Webb & Palinscar, 1996). The peer-learning focus in this study is dyadic collaborative peer learning in a classroom context.

### *Questioning*

Another method for promoting conceptual-knowledge learning that is championed by educators and researchers is questioning (e.g., Chin & Brown, 2000; Graesser & Person, 1994; Hmelo & Day, 1999; King, 1989, 1994). A question can be defined as an inquiry (i.e., request for information) or interrogative expression (i.e., statement followed by a question mark), or both (Graesser & Person, 1994). Questions in an educational situation can include questions directed to teachers or peers from students, questions directed to students from teachers, or questions directed to students from a source such as a textbook. The latter two situations, questioning of students by teachers and other sources, have been common teaching methods for centuries (e.g., the Socratic method) and are the form of questioning that is a focus of this study.

Several approaches to questioning of students have been widely researched, particularly in the field of reading comprehension. Those approaches include adjunct

questioning, advance questioning, and elaborative interrogation (e.g., Anderson & Biddle, 1975; Hamaker, 1986; Pressley, McDaniel, Turnure, Wood, & Ahmad, 1987).

Adjunct questions are questions added to instructional text that are intended to influence what the reader learns from the text (Hamaker, 1986). These questions can be provided before relevant text as pre-questions (either massed at the beginning or inserted before each relevant passage), or as post-questions after the relevant text (either massed at the end or after each relevant passage). Andre (1979) identifies the primary types of adjunct questions that have been investigated in research. They include *factual questions*, which use verbatim language from a text a learner has read; *paraphrased questions*, which are factual questions with only minor word overlap from the text; *general questions*, which are factual questions that ask about more than one text sentence; *application questions*, which ask students to choose a novel example of a concept they have read about from a list of alternatives (selected-response); *meaningful learning or inference questions* which asks students to make connections between passages in a text that are not explicit; and *higher-order, analysis, or evaluation questions*, which have not been clearly defined in the literature but which usually are described as accessing categories of Bloom's (1956) Taxonomy at levels higher than knowledge (e.g., comprehension/understanding, application, evaluation).

Advance, or orienting, questions are adjunct pre-questions that learners receive prior to engaging in a learning task and that are intended to guide a student while they learn, similar to an advance organizer (Anderson & Biddle, 1975; Osman & Hannafin, 2001). They can take any of the forms listed previously, from factual through higher-order questions.

In elaborative interrogation, “why” questions about particular sentences in a text are asked of learners after the text is read (Pressley et al., 1987). Often these questions ask students to evaluate just-read factual statements with prior knowledge (e.g., Martin & Pressley, 1991). These include questions such as “why is this true/not true;” “why does this make sense based on what you know;” and “why did that particular man do that?” (Pressley et al., 1987; Pressley, Symons, McDaniel, Snyder, & Tunure, 1988).

Although these approaches employ distinct labels for types of questions, they do not represent mutually exclusive categories (e.g., advance questions are a type of adjunct pre-question, and adjunct questions could include elaborative interrogation “why” questions). Furthermore, at its core, each approach supports the idea that questioning a learner may facilitate active construction and integration of knowledge into the learner’s already-existing knowledge, and thereby facilitate the learner’s understanding of the material they are learning (Brown, Bransford, Ferrara, & Campione, 1983; Kintsch, 1998).

In another similar line of questioning research, Graesser and colleagues (Graesser & Person, 1994; Graesser, Person & Huber, 1992) have developed a taxonomy of questions that they derived from student and tutor-generated questions during one-on-one tutoring for research methods and algebra. Their analysis identified several types of questions that were associated with higher academic outcomes, what they call “*deep-reasoning*” questions. The questions are called “reasoning” questions because they elicit explanatory reasoning via logical and causal chains, which is hypothesized to foster deeper processing of information (Graesser et al., 1992; Graesser & Person, 1994; Hmelo & Day, 1999). These question types include *antecedent questions* (“what were the

reasons this experiment failed?”), *causal consequence questions* (“what happens when this level decreases?”), and *expectational questions* (“why isn’t there an interaction?”). Many questions classified as reasoning questions under the Graesser and colleagues scheme could also be classified as application, inference, or higher-order questions as explicated by Andre (1979). Graesser and colleagues use the term “deep” to describe the type of reasoning these questions elicit. However, the depth of reasoning required by a question is likely dependent on the wording and context of the particular question, and the Graesser et al. (1992) taxonomy does not account for this. As such, for the purposes of this study, these question types, classified under the Graesser et al. (1992) taxonomy, will be referred to simply as *reasoning questions*.

Despite evidence that these reasoning questions are associated with higher learning outcomes in a tutoring environment, the mechanism behind how these types of questions can foster conceptual-knowledge learning is still hypothetical and largely untested. This study investigated these mechanisms through analysis of students’ question responses, in conjunction with their conceptual-knowledge learning gains and with their collaborative discourse as they worked together to learn and to answer a reasoning question.

The next section presents the theoretical framework that guided this investigation of peer collaboration and questioning.

## Theoretical Framework

### *Theoretical Perspectives on Peer Learning*

The methods of implementing peer learning in the classroom are diverse. This is primarily because of the wide variety of theoretical perspectives that drive research and practice on peer learning. The most common theoretical perspectives include social-motivational (e.g., STAD; Slavin, 1986; CIRC; Stevens, Madden, Slavin, & Farnish, 1987), social-cohesion (e.g., Learning Together; Johnson & Johnson, 1991), sociocultural (e.g., CSILE, Scardemalia, Bereiter, & Lamon, 1994), cognitive-elaboration (e.g., O'Donnell, 1996; Palinscar & Herrenkohl, 2002; Webb, 1989), and Piagetian and Vygotskian theories (e.g., Delisi, 2002; Hogan & Tudge, 1999). Social-motivational and social-cohesion approaches emphasize the motivational aspects of peer learning more than cognitive processes involved in peer learning (e.g., Slavin, 1986). In contrast, cognitive-elaboration and the developmental Piagetian and Vygotskian theories focus more on the cognitive and social-cognitive processes of peer learning with less emphasis on motivational aspects (e.g., O'Donnell et al., 1990).

This diversity in theoretical approach has several implications for researchers and teachers interested in investigating peer learning. First, inherent in each theoretical perspective is a hypothesis about the mechanism behind *how* collaboration with a peer enhances learning. For example, social-motivational theories posit that rewards provided to the group as a whole will motivate each member to work hard at the task, which leads to more effective learning (Slavin, 1996), while the cognitive-elaboration perspective holds that students learn in collaborative situations through heightened use of processing activities (O'Donnell, 1996; Webb, 1989, 1991, 1992).

Second, the theoretical perspective drives the way in which students are grouped, the directions they receive, and the goal and reward structure provided while they collaborate (O'Donnell, 2006; Webb & Palinscar, 1996). For example, social-motivational methods would include group rewards (e.g., STAD, Slavin, 1986; CIRC, Stevens et al., 1987), whereas other approaches would not include any type of reward structure (e.g., Reciprocal Peer Tutoring; Fantuzzo, King & Heller, 1992; Scripted Cooperation; O'Donnell & Dansereau, 1992; unscripted collaborative learning; Webb, 1989, 1991, 1992). The result is a variety of different approaches for implementing peer learning in the classroom.

Third, much of the data from peer learning research come from collecting discourse between peers (e.g., Chinn, O'Donnell, & Jinks, 2000; Webb, 1991). However, the theoretical perspective from which the research comes necessarily shapes the choice of analytical lens for the data. For example, using a cognitive-elaborative framework, discourse would be analyzed for use of particular processing behaviors, but not necessarily any motivation-related processes. As such, there are multiple ways in which to view the same interaction or discourse depending on one's theoretical perspective.

In an effort to bridge various aspects of these theories, I have adapted the conceptual lens of self-regulated learning (SRL) to explore the dynamics of peer learning. Self-regulation is the process by which a person regulates his or her cognitive, emotional, or physical behaviors based on internal drives or motivations (Bandura, 1986; Schunk, 2001; Zimmerman, 2000). A particular aspect of self-regulation, self-regulated learning is included and modeled in many learning theories (e.g., information-processing theory,

Winne, 2001; social-cognitive theory, Schunk, 2001) to explain the processes by which students learn.

I drew upon several theoretical and conceptual bases related to SRL to begin looking at students' regulation of learning in a collaborative context. This multifaceted approach is necessary because none of the various theoretical models address SRL in a collaborative context specifically, although some researchers have begun to address the role of collaboration in the development of SRL at a conceptual level (e.g., McCaslin & Hickey, 2001). As such, it is necessary to draw from several theoretical sources to ground the current investigation into *collaborative* self-regulated learning. Social -cognitive theory (Bandura, 1986) is the overarching theoretical framework that guided this study because of the emphasis placed on the interplay between context (environment) and person, as well as the central importance of self-regulation in the theory. However, despite using this theory as an overarching one, I also drew from aspects of other theoretical approaches and relevant research to explore the specific processes involved in collaborative learning.

### *Social-Cognitive Theory*

Social-cognitive theory seeks to explain and predict human behavior and its development. The theory, as described by Bandura (1986), is based on several key ideas. First, according to social-cognitive theory, behavior is explained using a model of interaction between people's environment, their behavior, and personal factors such as cognition. These components influence and affect each other in "triadic reciprocity" to determine human functioning (Bandura, 1978, 1986). As such, social-cognitive theory

recognizes the importance of the interplay between the self, the environment, and behavior in shaping human thought and development.

The second key idea is that people are assumed to have a number of capabilities that influence their development. These include the ability to use symbols, the ability to engage in forethought, the ability to learn vicariously by observing others who model particular behaviors, and the ability to regulate oneself based on internal standards and evaluations (Bandura, 1986). These capabilities allow humans to communicate, interact with their environment, and learn from themselves and others. Inherent in both these assumptions is the idea of agency (Bandura, 1986). Specifically, people are thought to actively interact with their environment—influencing it as well as being influenced and affected by it (Bandura, 1986; Zimmerman, 2000).

Overall, social-cognitive theory provides a viable theory for addressing many aspects of human behavior. The next sections address how social-cognitive theory and its assumptions can guide research on peer learning in the classroom.

### *Social-Cognitive Theory and Peer Learning*

Peer learning fits within a social-cognitive perspective on human learning in two primary ways. First, peer learning can be viewed from a social-modeling perspective. Observational, or vicarious, learning is a core assumption of social-cognitive theory, as it explains how people can learn skills and behaviors without actually performing them (Bandura, 1986). This is inherently a social component of the theory, because people must observe other people to learn. From an observational learning (or modeling) perspective, the mechanism behind peer learning is based on students learning skills and behaviors by observing each other as they work together (O'Donnell, 2006). Research on

peers as models has demonstrated that students will model peers if they feel that peer is competent at the task and if the task seems appropriate for the context (Schunk, 1987).

For conceptual-knowledge learning in particular, one can imagine that students may observe a peer engaging in behaviors that enhance their ability to learn challenging information, and then they subsequently use those observed behaviors in a similar situation. These behaviors could include enacting strategies such as making inferences or summarizing a difficult paragraph after reading it. Modeling of such strategies and behaviors by teachers has been recognized as effective in a variety of domains such as writing (e.g., Schunk & Swartz, 1993). Indeed, the modeling perspective is the dominant one taken on school-based learning for social-cognitive researchers (e.g., Schunk, 2001).

The second way in which social-cognitive theory may explain and predict outcomes of collaborative learning is through its assumption of the regulatory capability of people as they learn. It is this aspect of the theory that informed this research study.

*Self-regulation.* Self-regulation is a major component of the social-cognitive theoretical perspective. In social-cognitive theory, self-regulation is assumed to have three interacting subprocesses: self-observation, self-judgment, and self-reaction (Bandura, 1986; Schunk, 2001). Self-observation is the process by which a person takes note of his or her own behavior; self-judgment is the process of comparing one's progress with goals or standards one hopes to achieve; and self-reaction includes the affective results of self-judgment. Each of these processes impacts the other and influences a person's motivation and self-efficacy for the task (Bandura, 1986). During self-regulation, the triadic reciprocity of personal factors, behavior, and the environment

cause a cyclical process of regulation to accommodate changes in each of these facets (Schunk, 2001).

Zimmerman (1998, 2000) presents a three-phase model of self-regulated learning based on social-cognitive theory of self-regulation. It begins with a *forethought* phase, during which time a learner sets goals and plans strategies for achieving the goals. At this stage, a person's motivational beliefs come into play, which influence the level of goal setting and planning (Zimmerman, 2000). The second phase is the *performance or volitional control* phase. This phase involves self-control and self-observation processes, as well as enacting of task strategies. The third phase is *self-reflection*. This phase involves judging one's performance in comparison to some internal standard or goal (Zimmerman, 2000). This is the phase in which a person monitors his or her understanding or progress toward the goals he or she has set. Most theories of self-regulation describe these processes as cyclical and recursive, rather than necessarily linear (Pintrich, 2000).

The phases of SRL are influenced by changes in the environment in which a person is learning (Bandura, 1986; Schunk, 2001). From a social-cognitive perspective, peer learning situations offer a context in which self-regulated learning is enacted, possibly collaboratively, with peers. As partners work together to learn or perform a task, they will be regulating their learning according to this particular context (Alexander, 1995). For example, they may use their partner for support by seeking help from them; alternatively, they may elaborate on a point when their partner seeks help from them, or utilize a strategy that helps both learners (e.g., Winters & Azevedo, 2005). The way a person regulates his or her learning in a collaborative context is likely to be different than

if he or she were engaged in a learning task alone. Collaborative regulation may serve as a mechanism for enhancing learning of people involved in collaborative-learning situations.

Social-cognitive theory encompasses all of human behavior, not just learning. As such, Bandura has little to say about cognitive processes involved in learning in particular. And, while a model of SRL has been developed under a social-cognitive perspective (Schunk, 2001), it is discussed at a conceptual level. As such, I drew upon other models and perspective of SRL to guide this study.

#### *Other Models and Perspectives*

In an effort to flesh out more deeply the particular processes involved in the regulation of learning, I drew primarily on two related sources for the current perspective on regulatory learning processes used in this study. The first was a framework developed by Pintrich (2000), which is intended to synthesize the many similar features and assumptions from different theories of SRL, including IPT (Winne, 2001) and social-cognitive theory (Schunk, 2001). The second is empirical results obtained from several years of research into students' use of SRL behaviors when learning with hypermedia, based on their think-aloud protocols (Ericsson & Simon, 1993). These results, analyzed using the framework developed by Pintrich (2000), provide a lens for looking at the results of the current study. As Winne and Perry state "theory typically advances in reciprocal and recursive interaction with work to engineer measures related to theory (p.533)."

#### *The Pintrich (2000) Framework*

The Pintrich (2000) framework is a particularly powerful tool for research on SRL because it provides a taxonomy of SRL processes drawn from many theoretical perspectives on SRL. His framework is defined by four assumptions that the models all share: first, learners are active, in that they make decisions and engage in behavior to further their knowledge or understanding; second, students have the potential to regulate their learning; third, students are aware of some goal or criterion to which they should compare their progress; and fourth, the SRL activities mediate between the context and individual and the eventual achievement for that individual. The model designates four areas for regulation during learning: cognition (e.g., goal-setting, employing and monitoring of cognitive strategies); motivation (e.g., self-efficacy beliefs, values for the task, interest); behavior (e.g., help-seeking, maintenance and monitoring of effort, time use); and context (e.g., evaluation and monitoring of changing task conditions). It is assumed that students will cycle through the SRL phases of planning, monitoring, controlling, and reflecting in these four areas while they learn, however the degree to which this occurs depends on the context (Pintrich, 2000), for example when working with a partner versus alone.

Pintrich's (2000) framework necessarily remains at a particular level of abstraction. For example, under control of cognition, he lists "selection and adaptation of cognitive strategies for learning" without naming particular strategies. This level of abstraction is necessary because SRL is variable depending on the context. To hypothesize about the particular processes one might see in a collaborative context, I turned to prior research on collaborative learning from an SRL perspective.

*Research on Collaborative Self-Regulated Learning*

Regulatory processes are related to learning, and it is the quality and quantity of these processes that distinguished effective and ineffective forms of learning (Zimmerman, 2000). Further, because regulation of learning is context dependent, the processes a student engages in will change from situation to situation. The context for this study was a collaborative learning environment with a task that included, for some, answering a reasoning question about a challenging science topic, the circulatory system. What types of regulatory processes do students in this context engage in? In particular, what types of forethought, performance, and self-reflective activities do students use? Social-cognitive theory and other theories of SRL do not provide explanations at this level of analysis. As such, I drew from prior research about learners in collaborative learning situations to flesh out particular regulatory processes associated with this particular context.

Research conducted under a cognitive –elaboration theoretical perspective of peer learning adds insight into particular cognitive processes and strategies students may engage in while learning collaboratively. Based on information-processing theory, the cognitive-elaboration perspective holds that student learning may be enhanced in a collaborative context because their cognitive processing is heightened by the presence of a peer to learn with (O'Donnell & King, 1999; O'Donnell & O'Kelly, 1994). Research using this perspective has cited particular strategies that students use, such as elaboration, imagery, mnemonics, and providing explanations (e.g., Webb, 1989, 1991, 1992). In most cases, this research includes learning environments in which students are directed and trained to use such strategies with a peer (e.g., Scripted Cooperation, O'Donnell & Dansereau, 1992). While research using the cognitive-elaborative perspective can

contribute to the control of cognition aspect of regulation of learning, under the Pintrich (2000) framework, it has little to add to other areas and phases of SRL. Some prior research using the Pintrich framework in both individual and collaborative learning situations has begun to address these other areas of the SRL.

Prior research by Winters and Azevedo (2005) on collaborative learning in a similar context revealed a number of different regulatory processes that can be classified under the forethought, performance, and reflection categories of regulation. These processes were evidenced from verbal data collected while students worked collaboratively. For example, students in this context verbalized goals and plans, which are types of forethought behavior; they enacted performance strategies to help them learning in the particular context, such as summarizing information, taking notes, and asking a partner to clarify something they did not understand; they engaged in reflection to the extent that they monitored their progress toward their goals and monitored how well they were understanding what they learned (Winters & Azevedo, 2005). Further, the literature on questioning in learning contexts hypothesizes several particular strategies that are integral to answering reasoning questions effectively, such as engaging in explanatory reasoning (Graesser et al., 1992; Graesser & Person, 1994).

It is the aforementioned observed and hypothesized processes associated with peer learning and questioning that provided the foundation for the hypotheses and methods for this study. Analysis of students' discourse provided a window into the mechanisms behind the hypothesized role of peer collaboration and reasoning questions in fostering conceptual-knowledge learning. Further, the results of this study serve to expand the current literature on peer collaboration by using a theoretical perspective of

collaboratively-regulated learning through which to analyze the mechanisms driving students' collaborative learning. Results of this study also expand current models of self-regulation to include collaborative regulation, as it occurs between peers working together. By using a variety of approaches, including a social-cognitive model of self-regulation, Pintrich's (2000) conceptualization of a framework for SRL, and empirical results, I included many of the constructs associated with peer learning that have traditionally been investigated in isolation using one particular theoretical perspective. The result is not only an addition to SRL theory and research methodology from a collaboratively-regulated learning perspective, but a way to capture aspects of different theories of peer learning together under one lens, not just cognitive elaboration, but also motivational variables and processes. This study extends current theories explaining collaborative peer learning and provides a potentially more comprehensive theoretical and methodological approach to investigating peer learning through exploration of students' regulatory processes as they learn collaboratively.

The next two sections present the context for the study. In this study, students used a computer-based hypermedia encyclopedia to learn about the circulatory system, a complex biological system. As such, the first section describes complex systems, which is the category of conceptual knowledge students were asked to learn. The subsequent section describes computer-based learning environments (CBLEs) as one vehicle for learning about complex systems. The last section summarizes the study, presents the research questions and hypotheses, and provides operational definitions for the key constructs in this study.

### Contextual Framework

### *Complex Systems*

According to the AAAS (1993) and NRC (1996), much of the conceptual knowledge high-school students should be learning includes knowledge about complex systems, such as evolution, chemical equilibrium, and human body systems. Developing conceptual knowledge about complex systems, in particular, involves learning declarative knowledge about the structures of the system (i.e., their form) and their behavior, as well as understanding their interactive functioning (Chi & Ohlsson, 2005; Collins & Ferguson, 1993; Hegarty, 2005; Hmelo, Holton, & Kolodner, 2000; Hmelo-Silver & Pfeffer, 2004).

When the knowledge of these elements and processes are integrated into a conceptual model or schema, the knowledge is considered conceptual knowledge (e.g., schemata, mental models; Alexander, Schallert & Hare, 1991; Schraw, 2006). For example, conceptual knowledge of the circulatory system includes knowledge about the capillaries throughout the body. The structure is the capillary itself. Its behavior consists of allowing gasses to pass through its walls because of its thinness; the interactive functioning of the capillaries is to allow the blood being pumped by the heart to deposit and gather gases and nutrients and waste molecules in appropriate organs in the body.

As students gain conceptual knowledge, the knowledge becomes more principled, in that students are more easily able to integrate and maintain new knowledge and to apply it to novel situations (Alexander, 2003; Bransford, Brown, & Cocking, 2000). Students' conceptual knowledge about complex systems is important as these systems become increasingly integral to many science and non-science professions (Jacobson & Wilensky, 2006; Sabelli, 2006). As such, conceptual knowledge of complex systems is

critical for today's high-school students who may need to build upon or apply this knowledge in a future professional context.

*Computer-Based Learning Environments, Multimedia, and Hypermedia*

The use of computer-based learning environments (CBLEs) in science classrooms has been touted as a means through which learning of conceptual knowledge of complex systems can be supported (Jacobson & Wilensky, 2006; Lajoie & Azevedo, 2006).

CBLEs are defined as technology environments that give students access to different representations of scientific data and information, such as text, diagrams, graphs and manipulable models (Derry & Lajoie, 1993; Gredler, 2004; Jacobson & Kozma, 2000; Lajoie, 2000). Some common features of CBLEs include multiple representations of information, non-linear access to that information, and some level of learner control in accessing the information (Jonassen & Reeves, 1996, Williams, 1996).

Because of the common feature of multiple representations of information within CBLEs, most can be classified as a type of multimedia learning environment (Mayer, 2005). Broadly, multimedia learning environments are computer environments in which images (either static or dynamic) and words (either auditory or visual) are presented together to convey information about a particular topic or concept (Mayer, 2001, 2005; Reed, 2006). The type of multimedia learning environment that is the focus of the current study is hypermedia. Hypermedia is a particular type of multimedia in which nodes or chunks of information in the form of text, diagrams, audio and video are linked so that the user can navigate between them easily (Dillon & Gabbard, 1998; Dillon & Josbt, 2005).

While CBLEs such as hypermedia hold great potential as tools to help students learn, past research on the use of CBLEs has demonstrated that students often have learning difficulties when using such environments (Azevedo, 2005; de Jong & van Joolingen, 1998; Graesser, McNamara, & VanLehn, 2005; Lajoie & Azevedo, 2006). CBLEs, in which learners have some control over their learning, such as hypermedia, present many choices for the learner and can be overwhelming (Alexander, Kulikowich, & Jetton, 1994). For students to learn about complex science topics using such CBLEs, they need to be effective regulators of their learning. They need to assess the task, set appropriate goals, choose and use strategies and particular science skills to meet those goals, and monitor their progress and understanding as they learn as they engage in scientific thinking and reasoning to acquire conceptual knowledge (Azevedo, 2005; Schraw, 2006; Schunk, 2005).

### *Present Study*

The purpose of this study was to examine if and how peer collaboration and reasoning questions affect high-school students' conceptual-knowledge learning through analysis of students' collaborative regulatory processes while they use a hypermedia CBLE to learn about the human circulatory system. High-school students were the focus of this investigation, because students at this academic stage are often exposed to complex systems in science (AAAS, 1993; NRC, 1997) and, therefore, support for learning about these topics at this developmental level is important.

The research design was a mixed-method design, including both quantitative and qualitative analyses. The quantitative portion of the design was a 2 (Learning condition: Peer collaboration vs. Individual learning) X 2 (Questioning condition: Reasoning

question vs. No question) factorial design, with a conceptual-knowledge posttest measure as the dependent variable. Students were randomly assigned to work with a partner or to work alone, and dyads and individuals were then randomly assigned to answer a reasoning question pertaining to the circulatory system or to have no question provided to them. All students took a pretest to assess their prior conceptual knowledge about the topic, and this was used as a covariate in quantitative analyses.

All students used Microsoft Encarta, a hypermedia encyclopedia, to learn about the human circulatory system. Their overall task was to learn all that they could in 30 minutes about the topic using Encarta. Peers and individuals in the reasoning question condition also had the task of answering the provided question as they learned. This question, drawn from prior research on the role of questioning in human tutoring (Graesser, 1993; Graesser et al., 1992; Graesser & Person, 1994), asked students to synthesize and apply the information they were learning.

A subset of peer groups in both conditions was audiotaped as they learned collaboratively. After the 30-minute learning session, all students took a posttest to assess their conceptual knowledge after the task. Data analysis included comparing learning differences between students working individually and those working in groups, as well as between those provided a reasoning question and those without. Student discourse was analyzed qualitatively for evidence of collaborative self-regulatory processes in the planning, control, monitoring, and motivation phases conceptualized in the Pintrich (2000) framework of SRL and the social-cognitive model of self-regulation (Zimmerman, 2000).

### *Research Questions and Hypotheses*

The intent of this study was to help explicate further the role peer collaboration and reasoning questions play in conceptual-knowledge learning about complex topics with hypermedia by focusing not only the product of student learning, but also on discourse of students as they learned, using a theoretical model of collaborative self-regulated learning. As such, the research questions and hypotheses for the present study were as follows:

*Research Question 1: Does working with a peer facilitate a students' conceptual-knowledge learning about a complex science topic to a greater degree than learning alone?* The hypothesis for this question was that students working in the peer-collaboration condition would demonstrate more conceptual-knowledge learning as evidenced by significantly greater gains on the conceptual-knowledge measure from pretest to posttest compared to students working individually. The support for this hypothesis comes from the peer collaboration literature that has demonstrated that peer learning often facilitates learning to a greater degree than learning alone (see O'Donnell, 2006; Webb & Palinscar, 1996).

*Research Question 2: a) Do peers engage in collaborative self-regulatory processes as they learn together? b) Is there qualitative evidence that these processes are related to learning outcomes?* The hypothesis for this question was that students would engage in collaborative self-regulatory processes (i.e., forethought, performance, and reflection), and that there would be qualitative evidence that higher quality and quantity collaborative self-regulatory processes are related to greater learning outcomes. This hypothesis is based on social-cognitive theory's model of self-regulation, which supports

the idea that high-quality and quantity self-regulatory processes are related to positive learning outcomes (Zimmerman, 2000).

*Research Question 3: Does answering a reasoning question while learning about a complex science topic facilitate high-school students' conceptual-knowledge learning?*

It was hypothesized that students in the reasoning question condition would demonstrate greater conceptual-knowledge learning in the form of significantly greater gains on the conceptual-knowledge measure from pretest to posttest than those who were not exposed to such questions. The support for this hypothesis is derived from the questioning literature, which has shown evidence of the positive relationship between higher-order, reasoning questions and learning (Graesser et al., 1992; Graesser & Person, 1994).

*Research Question 4: Does working with a peer facilitate explanatory reasoning associated with answering a reasoning question to a greater degree than working alone?*

It was hypothesized that students working with a peer would provide answers to the reasoning question that included a higher quality and quantity of causal chains, as evidence of their explanatory reasoning. The support for this hypothesis comes from the literature on the role of questioning in tutoring, and the observation that explanations are built collaboratively in one-on-one tutoring situations (Graesser, Person, & Magliano, 1995).

*Research Question 5: a) Do collaborative dyads answering a reasoning question while learning engage in greater instances of utterances related to forethought, performance, reflection, or motivation than collaborative dyads not answering a reasoning question?* For this research question, it was hypothesized that peers who receive a reasoning question would engage in more collaborative self-regulatory

processes related to deeper processing of information, as evidenced by a greater frequency of utterances related to performance and reflection compared to peers who did not receive a reasoning question. The support for this hypothesis is derived from the questioning and self-explanation literature, which has found evidence that high-quality self-explaining includes making inferences and knowledge-monitoring (Roy & Chi, 2005). Asking students a reasoning question in a collaborative context may spur them to engage in more explanatory reasoning as they work with their partner to answer the question. As such, in attempting to answer a reasoning question, a student may make inferences between the different pieces of information they have learned, or connect new information to prior knowledge (knowledge elaboration), as has also been seen in the research on high-quality self-explanation (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, de Leeuw, Chui, & LaVancher, 1994; Roy & Chi, 2005).

## Operational Definitions

The following terms are used throughout this dissertation. The operational definition, as it is conceived in this research study, is given for each.

*Collaborative self-regulated learning:* Collaborative self-regulated learning is a term to describe the regulatory behavior of learners working with at least one other person in a collaborative learning context. It is model of regulatory behavior based on social-cognitive theory, which identifies three cyclical phases of regulation: forethought, performance, and reflection that are influenced by the context in which the learning takes place (Zimmerman, 2000), and a synthesis of other theoretical models of SRL that includes planning, monitoring, control, and motivation (Pintrich, 2000).

*Conceptual knowledge:* Conceptual knowledge is defined as the interaction of one's declarative and procedural knowledge, and it demonstrates one's understanding or "knowing why" about a particular topic (Ryle, 1994). For complex systems, this includes understanding of the structures and processes and how they interact to contribute to the functioning of the system (Hmelo et al., 2000; Hmelo-Silver & Pfeffer, 2004).

*Reasoning questions:* Reasoning questions are questions that elicit explanatory reasoning, through logical justifications and causal chains, from those being questioned, and in so doing are hypothesized to facilitate understanding of the information (Graesser, 1993; Graesser et al., 1992; Graesser & Person, 1994).

*Hypermedia:* Hypermedia environments are computer environments in which multimedia information nodes in the form of pictures and words are linked together to facilitate navigation and access by the user (Dillon & Gabbard, 1998; Dillon & Jobst, 2005). Pictures can take the form of static or manipulated objects, animations, and

graphs; words can be either written (visual text) or spoken (e.g., auditory narration; Mayer, 2001, 2005; Reed, 2006).

*Peer collaboration:* Peer collaboration is defined as a peer-learning situation in which two, three, or four students working together, face-to-face, in classroom setting towards a mutual goal of learning from a particular task (O'Donnell, 2006). For this particular study, peer learners worked in pairs.

## CHAPTER 2: LITERATURE REVIEW

This chapter provides grounding for the present study through a review of research literature on peer collaboration and on questioning in educational contexts, as well as a review of theory and research on learning with hypermedia. This grounding starts with an overview of what is currently known about peer collaboration and questioning as pedagogical practices and their effects on learning. The overviews provide a context for more exhaustive reviews of studies that investigate these practices for fostering science learning in particular. Moreover, the reviews allow for a subsequent discussion of what more we need to know about these pedagogical practices for fostering science learning, and in so doing provide the rationale for the current study.

To frame this review, I approached the literature using four inquiries. The first inquiry is: *What do we know about the effect of peer collaboration on learning?* To address this inquiry, I will provide an overview of reviews on the various forms of peer collaboration. The subsequent section is guided by an inquiry about the subset of the peer collaboration literature most relevant to this study: *What do we know about the role of peer collaboration for fostering science learning, in particular?* In this section, I provide an exhaustive review of studies on peer collaboration for learning science.

In the third section, I turn to the literature on questioning. The inquiry guiding this section is: *What do we know about the role of questioning of students in learning?* To address this inquiry, I sample from reviews and the literature on different types of questioning intended to foster learning. This leads to the final inquiry about the questioning research most relevant to this study: *What do we know about questioning for*

*fostering science-learning, in particular?* For this section, I provide an exhaustive review of the relevant literature.

In this chapter, I summarize what the review of the literature reveals about learning with a peer and learning through questioning, with a particular focus on learning scientific concepts and skills. I describe what we still need to know and argue that this study is a step towards addressing these gaps in the research.

In the final section, I provide an overview of theories of learning with multimedia and hypermedia that have both guided and been developed from research on learning with these environments. I then provide an overview of what research has demonstrated about learning with hypermedia and discuss how these theories and research impact and guide this study.

### Peer Collaboration

#### *What Do We Know About The Effect of Peer Collaboration on Learning?*

For the purposes of this review, peer collaboration is defined as peer-learning situations in which students work together, face-to-face, in a classroom context on a mutual learning task with minimal intervention from a teacher (Cohen, 1994; Webb & Palinscar, 1996). Cooperative learning (Slavin, 1980) can be considered a type of collaborative learning that utilizes structured interaction between members of the collaborative group during the mutual learning task. *Peer tutoring* will not be considered in this review, as it is generally regarded as distinct from collaborative learning (O'Donnell, 2006). Similarly, *reciprocal teaching* involves a teacher scaffolding, then fading, the use of self-questioning, summarizing, predicting, and clarifying in reading comprehension (Brown & Palinscar, 1989; Hacker & Tenent, 2002). Students using

reciprocal teaching work in groups, but this technique involves substantial interaction between teacher and students and as such, it is not considered peer collaboration as defined in this study.

To provide an overview on the research on collaborative learning, I draw from major reviews of research in this area, as well as seminal articles and chapters on the subject. A more exhaustive review of the literature specific to science learning is provided in the section that follows. However, this overview serves to situate the exhaustive review within larger context of what we know about the effects of peer collaboration on learning more generally.

The most effective way to structure a general discussion on collaborative learning is to use the theoretical perspectives that drive research in peer learning. O'Donnell (2006), O'Donnell and O'Kelly (1994), and Slavin (1996) provide comprehensive descriptions of the various theoretical perspectives that have governed research and use of peer collaboration in classrooms. These perspectives can be divided into two general categories: those that focus on social-behavioral aspects of learning, and those that focus on cognitive aspects of learning. While not mutually exclusive, these categories are based on the hypothesized mechanisms behind the success of collaborative learning (O'Donnell, 2006).

The social-behavioral perspectives are guided by motivational perspectives of learning (Slavin, 1996). One idea that is critical to this perspective is the idea of positive interdependence (Johnson & Johnson, 1991) in which individual success is dependent on the success of others in the group. This can be accomplished by providing different reward structures (e.g., STAD, Slavin, 1986), or by teaching students social skills and

structuring roles once they are in groups (e.g., Learning Together, Johnson & Johnson, 1991). While research has shown that these methods produce large effect sizes for performance of well-structured tasks, there is evidence that this is only the case if positive interdependence, in the form of individual accountability along with group goals and rewards, is present (Slavin, 1996).

The second group of theories focuses on the impact of peer learning on individual cognition. These include cognitive-elaboration perspectives and cognitive-developmental perspectives. Cognitive-elaboration perspectives hold that the advantage of peer learning over individual learning can be found in the enhancement to information processing activities that come with learning with a peer, who may help students monitor their understanding and help them stay on task (O'Donnell, 2006; O'Donnell & O'Kelly, 1994). These processing activities include schema activation, rehearsal, and metacognition. Examples of methods for this type of collaborative work include the structured Scripted Cooperation method (O'Donnell & Dansereau, 1992) in which peers are given particular tasks related to processing of text, such as summarization.

Cognitive developmental perspectives are grounded in Piagetian and Vygotskian theories. For example, those influenced by Piagetian theory see collaboration as a mechanism by which cognitive disequilibrium occurs, and in the process of reequilibration, the student constructs new understanding and knowledge (O'Donnell, 2006). Piagetian approaches often focus on the cognitive conflict that can occur when peers collaborate. Vygotskian perspectives hold that students can learn from each other when one is more knowledgeable than the other in a particular area (O'Donnell, 2006;

O'Donnell & O'Kelly, 1994). As such, approaches that espouse this perspective usually involve pairings of students of differing abilities.

The approaches employing social-behaviorist perspective tend to use groups of students (4-6) rather than dyads, and they tend to use the group structure for longer projects (e.g., weeks rather than days or hours). Conversely, cognitive-elaboration approaches group students in dyads because this is thought to maximize potential for interaction (Webb, 1989); the learning episodes under this perspective are often shorter learning episodes. Further, because of the hypothesized mechanisms driving research, the methodology of research studies utilizing the perspectives differs. For example, studies using a social-behaviorist perspective typically do not analyze student discourse. For cognitive-developmental researchers, however, the discourse between students is critical to finding evidence for hypothesized mechanisms.

What have reviews of research on collaborative learning found? Slavin (1983, 1996) has argued that there is considerable evidence that individual accountability with group rewards is the most effective method of collaborative learning, no matter what theoretical perspective one uses. In particular, he argues that one can explain any positive effects of cognitive elaboration, for example, through a motivational perspective. Essentially, students who are motivated to learn and to ensure that their group members learn also, will engage in processing strategies that are beneficial to learning (Slavin, 1996). He does describe three circumstances in which the individual accountability with group rewards is not necessary: when the task is controversial and no "right" answer exists; when the collaboration is in the form of a voluntary study group for an external assessment; and when the collaboration is a highly structured dyadic task (Slavin, 1996).

He calls for more research on conditions in which group rewards and accountability are not necessary to foster learning, particularly because many classroom teachers do not like to use rewards (O'Donnell, 2006; Slavin, 1996).

Cohen (1994) offers a different perspective in her review of the research on peer collaboration. She argues that the effectiveness of different collaborative techniques depends on the outcome variable of interest, what she terms "productivity" (p.3), as well as what type of task students are given. She distinguishes between different outcomes, the most common of which has been conventional academic achievement stressing basic skills and factual knowledge. Cohen (1994) identifies other outcome variables important to both research and teaching with peer collaboration methods; they include conceptual learning and higher-order thinking, equal-status interactions between peers, and desirable prosocial behavior. Cohen (1994) asserts that the task definition is as important as the outcome variables for helping determine when and what kinds of collaboration will be beneficial. For example, tasks with right answers and clear procedures (well-structured, Simon, 1973, 1978) will entail a much different type of collaboration than tasks that are ill-defined (ill-structured, Simon 1973, 1978) and focus on conceptual learning or process skills (Cohen, 1994).

For her review, Cohen (1994) primarily focuses on the role of interaction (in the form of discourse) between peers as they collaborate, and how this may be related to different task types and outcome variables. Her assessment of a handful of naturalistic peer-collaborations reveal that students do not automatically engage in high-level discourse when they work together, and that some encouragement—through instruction or otherwise—may be necessary (Cohen, 1994). Closer scrutiny of mixed results in research

relating frequency of interaction to learning outcomes leads Cohen (1994) to argue that sheer frequency of interaction is related to learning outcomes for ill-structured tasks, but that for well-structured tasks with correct answers, the quality of the interaction is more strongly related. She cites Webb's (1983, 1991) research that has shown that the act of providing (and in some cases, receiving) detailed explanations facilitates one's learning in a collaborative situation with more routine tasks.

In comparing studies that look at different levels of structuring student interaction (from highly constrained interaction to unconstrained interaction), Cohen (1994) concludes that when the objective of the task is learning for understanding, instructions and arrangements of the peer collaboration that structure or constrain interaction between peers will not be as effective as situations in which instructions foster maximum interaction and elaborated discussion. Providing students with roles such as facilitator, but not structuring the interaction beyond this, is one example of instructions that provide the latter type of interaction and discussion (Cohen, 1994). In short, Cohen (1994) concludes that for fostering conceptual understanding, allowing students full, unconstrained discussion is critical; however, she recommends some task instruction that encourages discussion and interaction is beneficial, such as setting problems or questions for discussion, or specifying roles.

In addition to task structure, a number of researchers have investigated the role of various student variables in peer collaboration, particularly within a Vygotskian framework. Vygotskian approaches focus on the physical and psychological tools that allow students to develop and learn with a peer, as well as the limitations that might be present due to differences in individual development (Hogan & Tudge, 1999). The idea

of individual development is critical to Vygotsky's concept of zone of proximal development, which only occurs when a child collaborates with another person (Vygotsky, 1978). In other words, social interaction is a necessary aspect of learning (Tudge & Rogoff, 1999). However, research has shown that not all social collaboration is beneficial to learning and that it can sometimes be harmful depending on factors such as ability (Lou, Abrami & Spence, 2000), gender (e.g., Tudge, 1989, 1992; Webb, 1991), interaction style (Forman & Cazden, 1994), and status (Cohen, 1994).

Under developmental approaches, the two most widely researched variables are ability and gender. Consistent with the Vygotskian perspective, research on the effect of ability grouping (e.g., low ability student paired with higher-ability student) has a consistent result demonstrating that students of low-ability profit from learning with a higher-ability peer, but not with one at their own level (Lou, et al., 2000; O'Donnell, 2006; Webb & Palinscar, 1996). High-ability students tend to do equally well in heterogeneous and homogeneous groups (Lou et al., 2000; O'Donnell, 2006; Webb & Palinscar, 1996).

Some results on the role gender plays in collaborative groups suggest that girls benefit most from groups that have equal numbers of boys and girls, or girls only (O'Donnell, 2006). However, Tudge (1992) found that after a problem-solving task, girls are more likely to regress after working in a same-gender pair than were boys who worked in a same gender pair. Furthermore, in groups that are either majority male or majority female, girls tend to defer to the males in the group, even though they may be in the minority (Webb, 1984; Webb & Palinscar, 1996). Social interaction styles also may play a role in the success of collaborative learning. For example, Forman and Cazden

(1994) describe 3 different patterns of interaction between 3 different collaborating pairs: parallel, associative, and cooperative. They describe qualitative evidence that pairs that were cooperative developed higher problem-solving skills during the task than those who displayed more parallel and associative patterns of interaction (Forman & Cazden, 1994).

In her research on the role of status in peer collaboration, Cohen (1994, 1984) has found that academic, peer, and social status is correlated with interaction when student work in a small group, and that interaction is a predictor of learning outcomes.

In sum, research on peer collaboration has demonstrated that there are many influences on learning in a collaborative context, including the amount and type of structure of peer interaction, the type of task provided, the ability gender, interaction style, and status of the students involved. As such, peer collaboration can take many different forms and have potentially different outcomes depending on the combination of factors involved.

The goal of this overview was to provide a backdrop and structure for reviewing the literature on peer collaboration in science more specifically. In the section that follows, I look more specifically at the existing research on peer collaboration for fostering science learning.

*What Do We Know About The Role of Peer Collaboration for Fostering Science Learning?*

The aim of this section is to provide an exhaustive review of the research on peer collaboration for learning science. This review provides support for the present study by determining whether and how peer collaboration can facilitate learning in science, as well as frame what we still need to know. The studies included in this review had to meet

several criteria directly relevant to the proposed study. First, as an important theoretical and methodological aspect of this study is the role that discourse plays in student learning as they collaborate, only studies that included analysis of learning outcomes as well as discourse were included. In other words, studies that investigated student learning outcomes only or discourse only were not included in the review (e.g., Herrenkohl & Guerra, 1998; Hogan, Nastasi, & Pressley, 1999; Kaartinen & Kumplainen, 2002; Van Boxtel & Roelofs, 2001). Second, the learning task and content had to be described fully enough to judge whether the task was well-structured or ill-structured. Third, the domain of learning had to be science. Because these criteria narrowed the number of studies considerably, all developmental ages were considered in the review. Nineteen studies fit the criteria, and they are reviewed in this section.

To organize this review, I draw on Cohen's (1994) assertion of the importance of task type in peer learning outcomes. In particular, I focus on well-structured tasks as compared to ill-structured tasks (Frederiksen, 1984; Simon, 1973, 1978). To focus the review on science learning in particular, I also use learning outcomes as a method of organizing and discussing the reviewed studies. Conceptual learning was a focus of this study, and as such, I further distinguish between studies that measure conceptual and factual knowledge and those that measure science process skills such as argumentation or reasoning, within each category. I do not separate factual and conceptual-knowledge learning in this review, as none of the reviewed studies distinguished between factual and conceptual knowledge. Further, none of the reviewed studies included enough information (e.g., copies of pretests and posttests) to allow for a judgment about type of knowledge assessed.

This structure allows for generalization from this research, particularly regarding the role of discourse which, Cohen (1994) argues, is influenced by task type. As such, the reviewed studies are divided into two groups: those investigating learning with well-structured tasks (e.g., worksheets, explicit instructions) and those investigating learning with ill-structured tasks (e.g., open-ended problem-solving, controversial topics). For science-learning, much of what students do in the classroom can be considered well-structured tasks. However, ill-structured tasks are considered more applicable to what occurs in real-life, and there is a call to use them in the science classroom, as can be seen with the emphasis on authentic inquiry in the science-education standards (AAAS, 1993; NRC, 1996).

It should be noted that many of the reviewed studies do not themselves focus heavily on the task type or knowledge type, but rather some other aspect of peer learning such as ability grouping. The purpose of organizing the studies in the manner described is to highlight features most pertinent to the proposed study: task structure and conceptual knowledge. Further, as suggested by Cohen (1994), these factors can greatly influence the role discourse plays in learning.

In reviewing the studies for level of structure for student interaction, only four of the 19 studies provided any well-defined and described structure for interaction, such as in cooperative learning techniques (i.e., Bianchini, 1997; Hogan, 1999; Kneser & Ploetzner, 2001; Saleh, Lazonder, & De Jong, 2005). As such, while this appears to be a critical factor in outcomes of peer learning according to Cohen (1994), it is not well-researched in the literature meeting the criteria for this review. This is likely due to the lack of interest in discourse in research on cooperative-learning methods.

*Studies with Well-Structured Tasks*

Well-structured tasks are defined as those that have clear procedures and “right answers” (Simon, 1973, 1978). The steps and procedures for completing a well-structured task are apparent and the criteria by which the answer is evaluated are clear (Simon, 1973, 1978). In the studies reviewed, six of the 19 studies investigated collaborative learning with well-structured tasks (i.e., Carter & Jones, 1994; Kneser & Ploetzner, 2001; Saleh et al., 2005; Tao, 1999; Tao & Gunstone, 1999; Webb, Nemar, Chizhik, & Sugrue, 1998). All of the reviewed studies that investigated well-structured tasks also used measures of conceptual and factual knowledge learning as outcome variables. None investigated scientific process or reasoning skills as an outcome measure.

Carter and Jones (1994), Saleh et al. (2005), Webb et al. (1998), and Kneser and Ploetzner (2001) investigated the role of prior ability or knowledge in collaborative group work. Using a Vygotskian perspective, Carter and Jones (1994) categorized middle-school students into one of three quartile ability groups: high, average, or low based on their scores on the California Achievement Test (CAT). Low and high-ability students were then paired with a student of the same or opposite ability. The pairs used a lab worksheet to investigate balancing with levers. Using a pretest and posttest measuring conceptual knowledge (factual and application questions), they found that low-ability students achieved more when working with high-ability peers, who did not differ in their ability no matter with whom they worked. These results support the findings of other ability-grouping studies (Lou et al., 2000).

The analysis of discourse in the Carter and Jones (1994) study focused on number of words spoken (frequency), number of speaking turns, block movement (task directed

activity with balance), tinkering (non-task directed activity with balance), helping behaviors, and distracting behaviors. The analysis of discourse revealed that overall, high-ability students spoke significantly more words and moved the block significantly more than low-ability students. However, low-ability students paired with a high-ability partner spoke significantly more words than their counterparts paired with other low-ability partners. High-ability partners of low-ability students also spoke more words and engaged in more helping behaviors than high-ability partners paired with other high-ability students. In contrast, low-ability students working with other low-ability students engaged in more tinkering and distracting behaviors, both considered negative learning behaviors in this context. In sum, low-ability students benefited from working with a more able partner, which may be related to greater interaction with their partner and less off-task behavior.

In one of the few studies in this review to look at highly-structured interaction, Saleh et al. (2005) used the cooperative learning technique called STAD (Students Teams Achievement Divisions; Slavin, 1980) to investigate elementary-school students' collaborative factual learning of botany in mixed-ability groups. Despite the structured nature of interaction in STAD, the results from this study are consistent with Carter and Jones's (1994) findings as well as those of other ability-pairing research (Lou et al., 2000); lower-ability students benefited from working with peers of higher-ability, and higher-ability students' achievement did not differ depending on the group in which they worked. Discourse analysis in this study focused on the type of elaborative talk that students engaged in. Saleh et al. (2005) found that homogeneous groupings, except those composed of low-ability students, fostered more collaborative elaborations—building on

each other's thoughts, answering questions and reasoning about content—compared to heterogeneous groups. Heterogeneous group dialogue yielded more individual elaborations, with low-ability students asking many of the questions (eight times as many as average-ability students), and high-ability students providing many (75%) of the explanations.

Webb et al. (1998) found similar results when middle-school students of differing abilities were grouped to work on an electrical circuit task together. Consonant with results from the previously reviewed studies, lower-ability students working in groups with at least one high-ability student had higher scores on a conceptual-knowledge assessment than did those working without a high-ability student. Contrary to the previous two studies, high-ability students working in homogeneous groups had higher achievement than those working in heterogeneous groups. Analysis of student discourse in the groups revealed that the groups with at least one high-ability students gave more correct answers and higher-quality explanations during the task than groups with lower-ability students. Further, students that were below average in the groups with higher-ability students were highly engaged in the discussions (e.g., asked questions, made and defended suggestions).

Kneser and Ploetzner (2001) conducted a study similar to the ability studies, but their focus was the role that prior knowledge type has on how students collaborate to solve classical mechanical physics problems. In this study, one group of high-school students was instructed on qualitative physics concepts and another group on quantitative physics concepts. A third group, receiving no instruction served as a control group. Dyads created by pairing students from the two experimental groups then collaboratively

worked on physics problems that required use of both qualitative and quantitative knowledge. Kneser and Ploetzner (2001) found no differences in conceptual-knowledge learning overall between students in the two instructional conditions, although they did show significant gains compared to the control group. However, an analysis on specific question type (i.e., qualitative or quantitative) revealed that students who were instructed qualitatively learned more from their quantitative partners than the other way around.

For the analysis of the discourse from these tasks, the researchers focused on dialogue structures, and specifically how the structures related to the roles students took during the problem solving and how well they learned. They compared three dyads that had differential learning outcomes: a pair in which both students improved from pretest to posttest; a pair in which only the qualitatively-instructed student improved; and a pair in which only the quantitatively-instructed student improved. The results showed that the most successful dyad had more coherent dialogues (as defined by average number of turns in a “dialogue game” –which is defined as a series of turns in which students pursue a goal and take roles accordingly); that the successful learners in each dyad sought information actively, either from their partners, themselves, or the environment; and that the successful learners often took the role of reflector, which manifested in more reflection occurring during the learning process.

In two studies investigating the role of co-construction of knowledge and conflict in collaborative learning for well-structured tasks in physics, Tao (1999) and Tao and Gunstone (1999) used a mix of Vygotskian (co-construction) and Piagetian (conflict) theoretical approaches. Tao (1999) compared dyads to individual students completing a qualitative physics test, during which students were asked to solve the problems and then

explain their answers. Similarly, Tao and Gunstone (1999) paired high-school students to work on a multimedia microworld to collaboratively predict, explain, and observe different physics experiments within the microworld.

Tao (1999) found that dyads performed better on the assessment than individual students, and analysis of discourse between pairs suggested that co-construction of knowledge occurred often, and that sometimes conflict would occur as well. Both often led to correct answers, lending support to both theoretical perspectives. In quite a different result, Tao and Gunstone (1999) found that fewer than half the 14 students demonstrated substantial conceptual change at posttest in their study, and half the students showed no change at all. Qualitative analysis of the discourse between pairs revealed that students who successfully changed their conceptions reflected on their conceptions and engaged in reconstructing them. While this was often a shared activity, Tao and Gunstone (1999) assert that this had to be a personal construction as well as a shared one. This result suggests that while a Vygotskian co-construction of knowledge may be beneficial, students must also develop an individual understanding to effectively learn when working collaboratively.

*Summary.* The studies that investigated ability and knowledge differences in collaborative groups in the science classroom provide evidence that lower-ability students' conceptual and factual-knowledge learning is facilitated when working with a higher-ability peer. Discourse in each study demonstrated that lower ability students tend to rely on higher-ability peers by asking more questions and engaging in greater discourse than when working with peers of similar ability. Higher-ability students working with lower-ability students tended to provide explanations to their peers.

However, these studies present mixed results on how this may affect high-ability students' learning. Carter and Jones (1994) and Saleh et al. (2005) did not find heterogeneous groupings hindered high-ability students' learning, but Webb et al. (1998) did. Tao (1999) provides support for peer collaboration as a positive pedagogical technique for physics assessments, while Tao and Gunstone (1999) provide evidence that peer collaboration may not always be successful. However, for this particular study, there was no comparison group, so we cannot be sure if individuals would have fared any better or worse than their collaborating peers.

As for the discourse analyses, we can begin to make some generalizations about potential mechanisms for successful peer collaboration with well-structured tasks, based on these studies. First, lower-ability students working with a higher-ability peer are aided through asking questions of their peer and greater on-task behavior than their counterparts working with another low-ability partner. Second, successful learners in collaborative activities, independent of ability, actively seek information and engage in reflective activities as they learn. So, while the collaborative environment may foster positive learning processes and provide resources for lower-ability students, it appears clear from the results of these studies that the individual is ultimately the deciding factor in whether learning occurs.

#### *Studies with Ill-Structured Tasks*

Ill-structured tasks are usually defined as task without defined procedures, and which have no one criteria or standard against which the outcome or answer can be evaluated (Simon, 1973, 1978). While many tasks in real-life may be of this nature, few classroom tasks fit this description wholly; instead, ill-structured classroom tasks usually

fall somewhere on a continuum between well-structured and the strict definition of ill-structured (Frederiksen, 1984). For this review, a task was considered ill-structured if the procedures were not all defined, and if the answer could not be considered merely correct or incorrect.

Twelve of the 19 studies reviewed focused on ill-structured tasks. The ill-structured tasks in the studies reviewed include: designing and conducting an experiment or reasoning about experimentation (Faulkner, Joiner, Littleton, Miell, & Thompson, 2000; Lumpe & Staver, 1995; Okada & Simon, 1997; Teasley, 1995; Winters & Azevedo, 2005); engaging in an open-ended joint task that results in a joint product such as a poster or presentation (Bianchini, 1997; Kumplainen, Solovaara, & Mutanen, 2001); and arguing and reasoning about scientific beliefs or statements (Alexopoulou & Driver, 1996; Chan, 2001; Chinn, O'Donnell, & Jinks, 2000; Hogan, 1999; Suthers & Hundhausen, 2003). While the majority of studies reviewed in this section investigate conceptual and factual learning as outcome measures, five of these studies (Alexopoulou & Driver, 1996; Chinn et al., 2000; Faulkner et al., 2000; Okada & Simon, 1997; Teasley, 1995) investigate scientific problem-solving, reasoning or argumentation skills as learning outcome measures. This difference in type of outcome measure will serve as a meaningful way of organizing these studies; it is possible that collaboration may differentially affect different types of knowledge-learning (Cohen, 1994). And, as conceptual-knowledge learning is the focus of this study, it is necessary to look the effects of collaboration on conceptual-knowledge learning in particular.

*Conceptual and factual knowledge learning.* Lumpe and Staver (1995) and Winters and Azevedo (2005) investigated student collaboration with scientific-

experimentation tasks. Working from a traditional Vygotskian perspective, Lumpe and Staver (1995) compared heterogeneous (by ability) triads to students working individually. They wanted to see what affect peer collaboration had on high-school students' conceptual understanding of photosynthesis after conducting experiments on plant growth. Their results showed that students working collaboratively had more correct conceptions than those who worked alone. However, Lumpe and Staver (1995) did not analyze students based on ability, merely on experimental group. Further, only the most successful triad's discourse was analyzed to find evidence for why group-work might facilitate learning. Students in this group engaged in sophisticated argument structures at various times, and that both consonant (friendly) and dissonant (conflict-oriented) exchanges occurred, and students took multiple roles during their interaction. The roles identified by the researchers for this group were: executive, skeptic, educator, record-keeper, and conciliator. These roles tended to fluctuate during the discourse based on students' expertise. From a Vygotskian perspective, the researchers assert that these role fluctuations occurred as a direct result of partners' perceived expertise, and that this enhanced concept development for the students in the group. However, absent any comparison with other groups, this remains speculative.

In a study focused on heterogeneous ability groups, Winters and Azevedo (2005) investigated high-school student pairs collaborating on experiments with a genetics microworld. Based on pretest scores, students were designated as low or high prior knowledge. Students were then paired in heterogeneous groups. Results demonstrated that the low-prior knowledge students made larger gains from pretest to posttest than the

high-prior knowledge students, although the high-prior knowledge students had higher posttest scores than the low-prior knowledge students.

Discourse analysis on the regulatory moves students made in these pairs demonstrated that low-prior knowledge students relied heavily on their high-prior knowledge partners for cognitive and regulatory support. In contrast, the high-prior knowledge students spent their time providing support for their partners or regulating their own learning. These results echo the results found by the ability researchers reviewed in the previous section. However, contrary to much prior research in ability groupings, this study used the Vygotskian perspective of co-regulated learning (McCaslin & Hickey, 2001), rather than co-construction of knowledge.

Bianchini (1997), and Kumplainen et al. (2001) investigated tasks that involved a joint product as the goal of the task. They found that students working with such tasks collaboratively spent much of their time on organizational and process activities. Bianchini (1997) investigated how middle-school students interacted in small groups while working collaboratively to create a presentation on the importance of blood. Students were paired in groups of 4-5 of mixed ability to work collaboratively on a presentation. The results from the study showed that gains from pretest to posttest across students were only modest; that despite the emphasis placed on concepts and connections, more time was spent by students on process and procedure; and that students with a high rate of talk had higher scores than students with a low rate of talk.

Kumplainen et al. (2001) found a similar result in their study. They had middle-school students work in dyads to produce a joint poster on an energy topic they researched together using a multimedia encyclopedia. Analysis of the discourse between

the groups revealed that the dyads spent more time on procedural, organizational, and management issues than on content. Further, evaluations of dyads' posters showed a lack of depth and coherence in conceptual understanding.

Chan (2001), Suthers and Hundhausen (2003), and Hogan (1999) all investigated how students worked collaboratively on problem-solving and reasoning tasks. All three used experimental or quasi-experimental designs to manipulate aspects of the task in an effort to determine the subsequent effect on collaborative activity, but with different foci in each study.

Chan (2001) investigated how task structure impacts how high-school students collaborate when learning about a controversial topic with incompatible information, such as evolution. Using a Piagetian perspective, she asked high-school students to rate the importance of a series of statements about evolution, which were organized based on whether students were in a knowledge-assimilation condition or a knowledge-conflict condition. Dyads in each condition were compared to singles in each condition. Findings revealed that peer collaboration did not result in greater conceptual change than individual learning in this situation. Further analysis of discourse patterns of successful and unsuccessful learners in dyads revealed that high-gain students engaged in more problem-centered moves (problem recognition, formulation of questions, and construction of explanations), whereas the low-gain group engaged in more surface moves (rating, ignoring, rejecting, and patching to eliminate differences) with their partners than high-gain students, independent of condition.

Suthers and Hundhausen (2003) manipulated the type of knowledge representation that collaborative peers used while engaged in a public-health problem-

solving task. College students used either graphical, matrix, or text notations to compile data, hypotheses, and evidence for the purpose of coming to a group conclusion about the cause of the public-health problem. These representational forms were chosen because they represent a hierarchy of necessity for identifying and constructing evidential relationships about the data (matrix > graph > text). As such, Suthers and Hundhausen (2003) predicted that pairs using matrix notation would talk more about evidential relations and elaborate more on previously represented information than those using graph and text notation, and that this would lead to greater conceptual learning.

Suthers and Hundhausen's predictions (2001) were supported to some extent. Matrix users did indeed talk more about evidential relations than graph and text users. Matrix and graph users engaged in greater elaborations on previous representation than did text users. However, while graph users included significantly more of the representations they created during the learning session in the post-assessment essay compared to the other two groups, there were no significant difference between the groups on a posttest measure of conceptual knowledge.

Hogan (1999) also found mixed results in a strategies-instruction study in which she trained half of the middle-school participant students on metacognitive, regulatory and strategic aspects of co-construction of knowledge. Students from each condition were placed in heterogeneous (on ability and gender) groups of four. These groups worked together to create and present their conception of a phenomenon they had observed. Analysis of the results revealed that even though students in the metacognitive training condition gained in metacognitive and collaborative knowledge (as evidenced by interviews), they were not able to put it into practice with peers on this task. She saw few

differences in collaborative and metacognitive activities while engaged in the task between the students who received the training and those who did not. Further, there were no significant differences between the groups on an application problem used as an outcome measure.

In sum, the research on peer collaboration with ill-structured tasks for learning conceptual scientific knowledge presents varied results. The role of ability in groups follows the previously reviewed research, in that high-ability students appear to provide support for lower-ability peers when they work together. The results also suggest that for some ill-structured tasks, students spend much of their collaborative time on procedural issues rather than on conceptual ones.

*Scientific skills learning.* Okada and Simon (1997) compared individual college students to pairs of college students working on a molecular genetics experimentation task using a computer microworld. Unlike most of the studies in this review that compared individuals to dyads, singles were asked to think-aloud while they worked to provide a comparison to the discourse pairs engaged in. Individual learners were compared to collaborative learners on the quality of their final hypotheses. This analysis revealed that pairs had better hypotheses than individual learners. Pairs were also more successful in their overall experimentation than were singles. Comparison of the discourse between the pairs with the think-aloud protocols from the individual learners showed that while pairs did not spend any more time or search the experiment space any differently than singles during the task, pairs participated in explanatory activities, such as entertaining hypotheses, talking about alternative ideas, and considering justifications more often than individual learners.

In a similar study, and using Vygotskian theory, Teasley (1995) compared individual learners to dyads carrying out a scientific reasoning task on the computer. Elementary-school students were placed into one of four groups: no-talk alone, talk-alone, no-talk dyad, and talk-dyad. Students' performance on the task was compared across groups. The results revealed that talk dyads were significantly more successful at the task than no talk alones and no-talk dyads.

Analysis of discourse revealed that the proportion of interpretive talk (e.g., planning, predicting, using strategies) was positively related to final hypothesis scores, whereas proportion of descriptive talk (e.g., procedural, describing evidence) was negatively related. Talk dyads engaged in more interpretive talk, whereas talk-alones engaged in more descriptive talk. The results of this study imply that talk, as compared to merely working with a peer (simulated by the no-talk dyads), may be the more important factor in collaborative learning situations, however, the results of this study clearly indicate that the type of talk that students engage in while collaborating, compared to learning alone, is important. The novel approach in this study, by including no-talk dyads and talk-alones, enables this type of assertion to be made. A potential methodological issue with this study and that of Okada and Simon (1997) is the fundamental difference between discourse and think-aloud data (Ericsson & Simon, 1993).

Using a socio-cultural perspective, Faulkner et al. (2000) investigated the role of ability grouping and task presentation in scientific reasoning with elementary-school students. Heterogeneous and homogeneous (by ability) groups of students were formed based on pretest scores from a scientific reasoning assessment. One half of the dyads worked on a chemical reasoning task on the computer, while the other half used a

physical, hands-on version of the task. Analysis of the data revealed that there were no significant differences in gains from pretest to posttest based on ability or environment type. However, analysis of gains from pretest to a delayed posttest revealed that same-ability pairs were more successful using the physical apparatus than using the computer, but no differences in task presentation were found with the mixed-ability pairs.

Analysis of student discourse focused on types of talk (task-related, social, procedural, and off-task) and on utterances related to reasoning (self-oriented or partner-oriented). These analyses revealed that students working in the computer condition had greater instances of task-related talk, social talk, and reasoning, and less procedural and off-task talk, than those in the physical apparatus condition. Further, mixed ability pairs talked more about the task and less about procedures than the same-ability pairs, but the same-ability pairs had a greater proportion of reasoning utterances. Based on these somewhat conflicting results between process and product data, Faulkner et al. (2000) concluded that reasoning utterances may not be related to the pretest and posttest assessment in this particular study.

Chinn et al. (2000) also investigated how the structure of a task impacted how middle-school students reasoned and argued during a collaborative discourse about electrical circuits. Groups of four were asked to collaboratively assess and discuss the quality of 3 statements about an electrical circuit. The groups were placed in one of two instructional conditions: to decide which statement was best and which was worst, or to decide which statements were “OK or not OK.” The best/worst condition was hypothesized to require more explanation, and reasoning between group members than the OK/not OK group. The researchers coded the groups’ arguments holistically, which

gave a score for complexity, as well as for specific features. The specific features analysis for the group consisted of variables such as number of nodes in the discussion and total number of collaborative-constructed arguments. As was predicted, the students in the best/worst condition had the more complex argument structures than those in the Ok/not OK condition.

Alexopoulo and Driver (1996) used a Piagetian perspective to research whether group size had any effect on how students reasoned about physics. High-school students worked in groups of either two or four to discuss physics reasoning questions. A sample of the discussions were recorded and then transcribed. The quantitative results suggest that students working in quads had more significant learning gains than those in pairs. Alexopoulou and Driver (1996) analyzed the discourse from progressive groups, in which one or more of the students gained from pretest to posttest, and regressive groups, in which the students either did not gain or showed a negative gain from pretest to posttest. The qualitative results were analyzed using a several levels of interaction. First, the researchers analyzed the argument construction of the students (e.g. prediction, justifications, evidence, and evaluating), then they looked at the social level of various statements (e.g. agree, disagree, asking questions), and finally, they classified the social dimension of statements (e.g. supportive, aggressive, uncertain). The researchers found that students in regressive groups spent a significant amount of time on social aspects, rather than actual argumentation. Students in progressive groups were open to their lack of understanding, and thus were hypothesized to learn more constructively from other group members.

*Summary.* The studies investigating ill-structured tasks in science provide mixed results to a greater degree than those investigating well-structured tasks. For conceptual-knowledge learning as well as for reasoning, task structure and presentation clearly affect the discourse that occurs between pairs (e.g., Chan, 2001; Chinn et al., 2000; Faulkner et al., 2000; Suthers and Hundhausen, 2003). However, these changes did not always foster greater learning. When working with open-ended joint-product tasks, students appear to spend too much time on organizational and procedural issues, to the possible detriment of learning (e.g., Bianchini, 1997; Kumpulainen et al., 2001). The few studies comparing peer learners to individual learners did find that peer collaboration in ill-structured tasks facilitated learning more than individual learning (e.g., Lumpe and Staver, 1995; Okada and Simon, 1997; Teasley, 1995), but not always (e.g., Chan, 2001). Finally, while regulatory activities seem to play a positive role in collaborative learning (e.g., Winters & Azevedo, 2005), training students to be more metacognitive and collaborative may not always help them do so in practice (e.g., Hogan, 1999).

With these varied results, it is more difficult to generalize about the role of student discourse in collaborative learning with ill-structured tasks in science. There is support for the positive role that explanation plays in learning with a peer (Okada & Simon, 1997); there is support for the beneficial effects of engaging in reasoning with a partner (e.g., Faulkner et al., 2000); and there is evidence that being metacognitively aware of one's lack of understanding is beneficial when working in a group (e.g., Teasley, 1995). However, more research investigating these aspects of student dialogue is necessary for these to be more than tentative evidence.

*Peer Collaboration in Science: What do we Need to Know?*

Based on the previous review, what more do we need to know about the role of peer collaboration in learning science? I will focus on two aspects of peer collaboration that need to be addressed further in research on this topic. First, it is clear from the review that the way in which the task is structured affects the role of collaboration in learning. More research on the way in which ill-structured tasks influence peer collaboration and learning in science is warranted. The review of this research revealed mixed results with respect to the role of peer collaboration with ill-structured tasks. Moreover, these tasks are increasingly emphasized in science education (e.g., AAAS, 1993; NRC, 1996), and thus there is a need to determine how best to implement them in the science classroom.

Second, research should continue to investigate and analyze discourse associated with peer collaboration. Discourse, in concert with learning outcome measures, provides detailed and rich evidence for the ways in which student talk can contribute to learning with a peer. However, the variety of ways of analyzing these data makes generalization across studies virtually impossible. As such, future research should focus on ways of synthesizing across research perspectives. While not a complete solution, Bandura's (1986) social-cognitive theory on self-regulated learning may be one initial approach in this direction. In fact, few researchers have used this perspective to analyze peer collaboration, despite the recognized importance of self-regulation for individual learning (Schunk, 2001). Because self-regulated learning encompasses students' forethought, performance, reflection and motivation, it encompasses many of the processes of interest to a variety of researchers.

As such, this study investigated students' collaborative science learning with an ill-structured task. Process data was collected and analyzed through the lens of social-cognitive theory to provide a novel, and perhaps more comprehensive, perspective on the role of discourse in peer collaboration and learning.

### Questioning

#### *What Do We Know About The Role of Questioning in Learning?*

Questioning, both by students and of students, has been posited to support student learning in various contexts (e.g., Callender & McDaniel, 2007; Chin & Brown, 2000; Graesser & Person, 1994; Hmelo & Day, 1999; King, 1989, 1994).

Research on questioning by students has taken two approaches. One approach has been to investigate student-generated questions in naturalistic environments with no training or intervention provided to students (Chin & Brown, 2000; Chin, Brown & Bruce, 2002; Costa, Caldiera, Gallategui & Otero, 2000). Research in this area has revealed that the quality and depth of students' questions depends on their prior knowledge about the topic. In particular, students who have some prior knowledge ask more higher-level questions, which are defined as questions that ask for explanations, inferences, application, or integration of information, than those with no or low prior knowledge (Scardamalia & Bereiter, 1992). Furthermore, students ask the most questions when the text they are reading is well-matched to their prior knowledge (Miyake & Norman, 1979). Also, while student question-generation has been shown to be an effective strategy for reading comprehension, it may be mediated by the level of question posed by the student (Taboada & Guthrie, 2006; Wong, 1985). For example, Taboada

and Guthrie (2006) found that the level of generated questions had a positive association with the level of students' comprehension after reading a science text.

The other approach to research on student-generated questioning has been to provide students with training on how to ask good questions prior to engaging in a learning task, usually involving reading text. Training in question-asking has taken the form of modeling good questions for students, and then providing students with question stems or prompts while they engaged in a learning task (e.g., King, 1994; King & Rosenshine, 1993). A review by Rosenshine, Meister, and Chapman (1996) found that, in general, providing students with training and prompts for question-generation has a positive impact on their question-asking ability and their comprehension of the material they are learning. In particular, training students to ask questions that tap prior knowledge was more effective at fostering learning than merely training them to ask basic comprehension questions (King, 1994).

While the evidence from research suggests that student-generated questioning enhances comprehension and conceptual-knowledge learning, the mechanisms behind why this occurs is as yet unclear (Taboada & Guthrie, 2006). Hypothesized mechanisms include the explanation that generating questions results in active processing of the information being learned (Wittrock, 1981); that generating questions may serve to focus a student's attention on important aspects of the text, the "selective-attention" hypothesis (Andre, 1979); that generating questions helps students connect prior knowledge with the information they are learning (Miyake & Norman, 1979); and that the level of question help students build knowledge structures from the information they are learning (Taboada & Guthrie, 2006).

Questioning *of* students, by teachers and other sources, has also been widely researched, particularly in the field of reading comprehension and posits similar mechanisms for learning as those set forth with student-generated questions. Several theoretical approaches to questioning have guided this research. These approaches include advance questioning, elaborative interrogation, and adjunct questioning (e.g., Anderson & Biddle, 1975; Hamaker, 1986; Pressley et al., 1987). At their core, each approach supports the idea that questioning a learner may facilitate active construction and integration of knowledge into already existing knowledge, and thereby facilitate understanding of the material being learned (Brown et al., 1983). The hypothesized mechanisms behind how questions can support learning have been most thoroughly explored in these various approaches to questioning.

Advance, or “orienting,” questioning is designed to support a learner’s metacognitive activities while they learn (Anderson & Biddle, 1975; Osman & Hannafin, 1994; Pressley, Tanenbaum, & McDaniel, 1990). High-level orienting questions, which activate a learner’s prior knowledge and require integration of new information, have been found to enhance problem-solving skills, particularly when learners are told the intended purpose of the questions (Osman & Hannafin, 1994).

Elaborative interrogation is a method of aiding comprehension of confusing information by asking “why” questions of the reader, which encourages them to construct an understanding of what they have read. This method has been shown to enhance a learner’s ability to remember factual information (Martin & Pressley, 1991; Pressley et al., 1987; Pressley et al., 1992; Pressley, Symons, McDaniel, Snyder, & Turner, 1988). For example, in study with young children (grades 4-8), an elaborative interrogation

condition led to statistically significantly more fact recall than conditions in which there were no questions, or where elaborations were provided to the students (Wood, Pressley, & Winne, 1990). One possible explanation for this is that the process of constructing an answer (an explanation) to a question is what facilitates learning, particularly as students do not tend to do this automatically as they are learning (Pressley et al., 1992). However, research on the effect of elaborative interrogation has mostly focused on memory of factual information, and only a few studies have investigated and shown elaborative interrogation's positive effect on higher-order learning processes, such as the ability to make inferences (e.g., Ozgunor & Guthrie, 2004). More recent research also suggests that the success of elaborative interrogation may depend on the comprehension ability of the learner, with low comprehenders not benefiting as much from these types of questions as higher comprehenders (Callender & McDaniel, 2007).

Another category of questioning research is that of adjunct questions. Adjunct questions are questions that are added to a text to have an effect on how the reader learns from that text. They can take many forms (including advance questions, and elaborative interrogation questions). Hamaker's (1986) extensive review of adjunct questioning reveals several factors that influence the effect of adjunct questions: the cognitive level of the questions, the position in the text of the questions, and the type of test used to assess learning. The review finds that higher-order adjunct questions intended to foster conceptual understanding are usually better facilitators of learning than factual adjunct questions, but that the task and content being learned often dictate the type of adjunct question best used and the placement of those questions. Research has also found that

adjunct questions can help students learn particular content with visual information as well, such as diagrams of scientific processes (e.g., Holliday & McGuire, 1992).

In a more recent approach to questioning, Graesser and colleagues (Graesser, 1993; Graesser, Baggett, & Williams, 1996; Graesser et al., 1992; Graesser & Franklin, 1990; Graesser & Person, 1994) have developed a cognitive theory of question asking and question answering based on interactions between students and tutors. They have developed a coding scheme of questions, which they derived from student and tutor-generated questions during one-on-one tutoring (Graesser, et al., 1992; Graesser & Person, 1994). Their analysis identified several types of questions that were associated with higher academic outcomes, so called deep-reasoning questions. Deep-reasoning questions (Graesser et al., 1992; Graesser & Person, 1994; Hmelo & Day, 1999) are thought to foster deeper processing of information. In particular, Graesser et al. (1992) identify six types of reasoning questions because they elicit logical, causal or goal-oriented systems reasoning in answering them.

According to the Graesser and colleagues' coding scheme, there are several types of reasoning questions, including causal antecedent, causal consequence, expectational, and enablement questions. These question-types include "why, how, what-if, what-if-not, and what are the consequences," as reasoning questions because they elicit explanatory reasoning that can lead to deeper understanding of the material being learned (Graesser et al., 1996; Hmelo & Day, 1999; Pressley et al., 1992). The research on self-explanation lends some support to the potential of reasoning questions for fostering deeper understanding. Self-explanation research has demonstrated that learners who spontaneously self-explain, or who were prompted to self-explain, had higher learning

outcomes than those who did not (e.g., Chi et al., 1989; Chi et al., 1994). Reasoning questions could be considered a type of prompt for self-explanation, and as such, the research on self-explanation may provide some potential mechanisms for their effect on learning. These potential mechanisms are explored in greater depth in the next section.

The various conceptualizations of questions presented in this section highlight the overlap that occurs between them. Higher-order adjunct questions are questions provided to a learner that require some inferencing and connection of information within the text on the part of the learner to answer them (Hamaker, 1986). As such, the Graesser and colleagues' reasoning questions can also be classified as higher-order adjunct questions, because they, too, require inferences and connection of information. I discuss reasoning questions in greater depth in the next section, as they are a particular focus of the proposed study.

### *Reasoning Questions*

For a question to be considered a reasoning question, the answer must not be verbatim in the information source, but rather, the learner must make inferences, elaborations, and connect information with prior knowledge to craft a response. As such, reasoning questions are classified as such only to the degree that the text or source requires these deeper processing activities. What would be considered a reasoning question related to one text might simply be considered a recall question with a different text.

How can reasoning questions foster conceptual knowledge learning? One theory that was explored in the early literature on questioning, particularly in regard to adjunct questioning, is that of the directed (or selective)-attention effect (Andre, 1979). The

directed-attention effect is a theory that students learn more when asked questions about the content because they are directed to more information in the process of seeking the answer to the question than those who are not asked questions. However, more recent research on questioning has revealed that the mechanism may be more complex. For example, research in elaborative interrogation has demonstrated that students learn more when they try to answer a question than if they do not try to answer it, and that the quality of the answer, if they give one, is not highly correlated with outcome measures (Pressley et al., 1990; Pressley et al., 1992). In both circumstances, the student has been exposed to the question, and thus has been directed to pertinent content; however, it is only in the attempt to generate an answer that the students' learning is facilitated.

Craig, Sullins, Witherspoon and Gholson (2006) suggest that reasoning questions (as explicated by Graesser et al., 1992) may stimulate mental model production and activate prior knowledge, which enhances learning (deLeeuw & Chi, 2003; Kintsch, 1998; Pressley et al., 1988; Pressley et al., 1992). Similarly, Pressley et al. (1992) contend that the important step in question answering is the *attempt* to answer the question. It is the attempt that helps integrate the new information with prior knowledge. They further conclude that for questioning to be a useful learning tool, learners must be prompted to provide an explanation as an answer. Merely asking the questions is not enough.

Similarly, research has shown that self-explanation involves making inferences as well as filling in missing knowledge or correcting faulty knowledge in mental models (Chi & Ohlsson, 2005; DeLeeuw & Chi, 2003; Roy & Chi, 2005). The process of "fixing" knowledge inherently requires a monitoring process, or judgment of learning,

whereby the learner becomes aware that he or she does not have a full understanding.

Craig et al. (2006) suggest that in attempting to answer a question, learners may become aware that their understanding is not complete or that they may have prior knowledge that is incongruous with the information they are encountering. Although Craig et al. (2006) do not identify it as such, this is a form of metacognitive monitoring.

The evidence from questioning, self-explanation, and reading-comprehension research provide possible mechanisms for explaining the potentially positive effects of reasoning questions on learning about complex topics. Further, these mechanisms can fit within a self-regulated learning framework and with collaborative learning. For example, when a student reads a reasoning question, it should cause them to think about what they already know, and in so doing activate their prior knowledge. According to Kintsch's (1998) construction-integration model, integrating new information with prior knowledge is a necessary first step to understanding what one is learning.

Within the SRL framework, this prior-knowledge activation is considered a cognitive planning activity. When a student reads a reasoning question, they may be prompted to engage in some monitoring of what they already do or do not know (feeling-of-knowing), as well as a monitoring of what they currently understand about the topic (judgment-of-learning). These are metacognitive monitoring processes within the SRL framework. Indeed, research on self-explanation has demonstrated that "high-quality" self-explaining, which is more positively related to learning gains than "low-quality" self-explaining, includes making inferences and knowledge-monitoring (Roy & Chi, 2005).

Reasoning questions may have a similar effect in stimulating students to engage in similarly high-quality learning processes (i.e., inferences, elaborations, and

monitoring) as students who are prompted to self-explain. As such, in attempting to answer a reasoning question, a student may make inferences between the different pieces of information they have learned, or connect new information to prior knowledge (knowledge elaboration), as has also been seen in the research on high-quality self-explanation (Chi et al., 1989; Chi et al., 1994; Roy & Chi, 2005).

These processes should be evident in a collaborative environment in which two students' task is to answer the reasoning questions collaboratively. In such a situation, students may verbalize their prior knowledge, their monitoring activities, and their explanations for their partner. The research on peer collaboration has demonstrated that these types of discourse moves do occur when peers collaborate (e.g., Kneser & Ploetzner, 2001; Okada & Simon, 1997; Teasley, 1995; Winters & Azevedo, 2005)

From a collaborative regulatory perspective, reasoning questions may activate a learner's prior knowledge, which is a forethought activity; learners may monitor their understanding in the form of judgments of learning and feelings of knowing, which are reflective activities; and questions may elicit the processes of inference generation and knowledge elaboration, which are performance learning strategies. For the proposed study in particular, these processes may mediate collaborative learning when such questions are provided.

#### *What Do We Know About Questioning for Fostering Science-Learning?*

A search for research on questioning as a strategy to foster learning in science revealed a number of ways in which questioning is addressed in science education research. The majority of the literature from this search can be grouped into four categories. The first category focuses on the assessment of science learning through

questioning. The second category pertains to discussions of ways to question students, but with little research involved; textbooks make up a majority of this section of the literature. The third category includes studies investigating questioning in the specific context of scientific inquiry (e.g., posing a testable question) in science. The last category includes studies that have investigated the role of various types of questions in conceptual science learning, outside of an inquiry context. It is research that falls in the latter group that comprise the studies reviewed in this section, as they are most pertinent to the proposed study.

Holliday and colleagues (Holliday, 1981; Holliday, 1983; Holliday & Benson, 1991; Holliday & McGuire, 1992; Holliday, Whittaker, & Loose, 1984) have investigated various aspects of the role of adjunct questions in science learning. For example, Holliday (1981) found that when students were provided with only a partial set of textbook study questions pertaining to a flow diagram, they performed worse on a posttest that covered content from the complete set of questions compared to students who were provided the complete set or who had no questions. This result supports the theory of selective attention, in that students' attention in the questioning conditions were directed toward whatever content the questions covered, but not as well to other content. In a similar result, Holliday and McGuire (1992) found that students provided with content-specific questions (in this case questions related to temperature or to heat) performed better on questions pertaining to that specific content than to questions that did not.

In a study also looking at the selective-attention hypothesis, Holliday and Benson (1991) randomly assigned students to one of five groups: control; no questions; questions

on easy content; questions on difficult content; and questions on both easy and difficult content. The task asked students to study a science chart on seven vitamins, and if they were in a question condition, answer the adjunct post-question related to each vitamin. A posttest, broken down by easy and difficult content, asked students to recall what they learned about the 7 vitamins. The results of this study suggest that students who receive questions focused on easy content tend to do better on this portion of the posttest, whereas students who answered questions on difficult content had an easier time on this portion of the posttest. In essence, these results also supported the selective-attention hypothesis.

Research by Holliday and colleagues has also demonstrated that not all question types are effective at helping students learn. Holliday (1983) found that comprehension adjunct questions that provided strong hints about the answers (overprompting) were not as effective for student learning as questions that did not provide such hints. High-school students were placed into experimental groups based on question type: prompted question, no-prompt question, no-question, and control. All groups but the control studied a picture-word diagram on biogeochemical cycles. Analysis of results on a comprehension posttest revealed that students who were provided unprompted questions outperformed students who were provided prompted question, who in turn outperformed those in the control group. In this study, students exercised the minimal effort necessary to answer the questions. For the prompted group, this came at a loss of learning.

In a similar study, Holliday et al. (1984) investigated the hypothesis that verbatim questions (those that mimic words in the text directly) interfere with learning by taking students' attention away from actually comprehending information. They were

particularly interested in whether low verbal learners would be particularly affected by such questions. In their study they found that, indeed, low-verbal students provided with verbatim questions performed worse on a comprehension posttest than low-verbal students who did not receive such questions as they learned. High-verbal students did not differ in their scores whether they received questions or not.

Leonard and Lowery (1984) investigated the role of different types of questions interspersed throughout biology text. These questions types were: rhetorical, which did not require an answer from the student; factual, which were based on recognition or recall; hypothesizing, which asked the student to predict the outcome; and valuing, which asked the reader to make a judgment or explanation. College students were randomly assigned to groups with different question types. All students took the same recall posttest. Interestingly, Leonard and Lowery (1984) found that the frequent questions, no matter the type, appeared to inhibit learning compared to the no-question group, who scored consistently significantly better than all the question groups on an immediate and delayed posttest. The authors mention in their discussion that students often commented that the questions were distracting for them. Further, they recognize that their posttest was mainly one of factual recall, and that a test of conceptual understanding might have had different results.

Osman and Hannafin (2001) investigated the role of conceptual orienting questions (advance questions) on factual knowledge and problem solving skills related to reading a text passage on Mendelian genetics. They found that students who were provided with orienting questions outperformed students without questions on both factual and problem-solving portions of the posttest. Further, students who were provided

a rationale for the question (e.g., “they will help you understand how to predict the genetic composition of an individual”) along with the question, scored significantly higher than those who received questions only. Trends in the data suggested that orienting questions improved problem-solving outcomes in particular, as all groups had similar scores on factual knowledge.

Riley (1986) conducted a study to determine the effect of question type on elementary-school students’ comprehension of information presented in a mini-science lesson. Unlike the previously-reviewed studies, this study investigated teacher’s question-asking during this lesson. Based on Bloom’s Taxonomy, teachers were randomly assigned to provide comprehension, knowledge or a combination of both types of questions during the lesson. Results revealed that on a comprehension posttest, the students who received both types of questions scored significantly higher than student who received comprehension only or knowledge-only questions.

Wang and Andre (1991) investigated the role of application adjunct questions in conceptual change with a text on electrical circuits. One group of students was provided with application adjunct questions, and the other group not provided with the question. Wang and Andre (1991) were also interested in the effect that pretesting might have, so they gave half the students a pretest, and the other half of the students received no pretest. Indeed, they found an interaction effect such that students who were given the adjunct questions scored significantly better than students who were not provided questions only when no pretest was given. The researchers speculate that pretesting effects may have overridden the effects of the questioning in fostering conceptual change. This result

appears to support the selective-attention hypothesis, in that the pretest may have directed students' attention to the important parts of the text.

The results presented in this review indicated that several factors, beyond their presence or absence, determine whether questions will help students learn. First, it should be noted that the majority of the studies reviewed investigate the role of adjunct questions, with only one study investigating the role of questions asked by teachers. As such, these generalizations apply mostly to adjunct questions in science.

One result from these studies is clear: questions appear to direct students' attention to important information, providing support for the selective-attention theory. As such, these questions need to be carefully considered when used in an educational context. The questions should be written in concert with the learning objectives (and ostensibly, then, the outcome measure).

Related to this result, questions that encourage active interaction with the material being learned are beneficial to learning. As such, verbatim or overprompting questions are detrimental to learning, as they do not require the learner to expend much cognitive effort to answer them.

Finally, Osman and Hannafin's (2001) study provides compelling evidence that providing a metacognitive component, in this case a rationale for the question, also helps students learn. However, more studies of this nature would need to be conducted to build a body of evidence supporting this finding.

#### *Questioning in Science: What do we Need to Know?*

The results summarized above provide guidelines for using adjunct questions in the science classroom. However, the studies reviewed here are quite narrow in their foci.

Most of the studies reviewed focused on factual learning and comprehension rather than conceptual understanding. One reason for this is that the topics of choice were not complex in nature. Also, while several of the studies investigated conceptual adjunct questions, most would not be considered reasoning questions under the Graesser and colleagues definition (Graesser et al., 1992; Graesser & Person, 1994).

As has been discussed in the introduction, the focus of science education has been increasingly directed toward complex conceptual learning. As such, there is a need to investigate the role of questions that can foster this type of learning. Do questions such as reasoning questions foster this type of conceptual learning in science? This question has not been empirically tested. Further, do theories such as the selective-attention effect apply to these types of questions, or is there another mechanism at work? Indeed, even the results from the reviewed studies hint at another mechanism - that of some level of active cognitive interaction with the material being learned. As discussed previously, several hypothesized mechanisms for the effect of reasoning questions include the positive role of self-explanation and other reading-comprehension strategies (Chi et al., 1989; Chi et al., 1994; Kintsch, 1998; Roy & Chi, 2005).

Another aspect of questioning that needs to be investigated is its role in contexts used often in current science classrooms, such as collaborative learning. None of the reviewed studies focus on the role of questioning in a collaborative context. What mechanisms may explain the role of questioning as two students work together? In fact, none of the studies on peer collaboration focus exclusively on questioning of students working collaboratively.

There also are directions the literature on questioning needs to take methodologically. The majority of questioning studies discuss hypothesized mechanisms but do not collect any process data to substantiate them. Future research should incorporate process data, through think-aloud (Ericsson & Simon, 1993), interviews, or discourse analysis to provide another source of evidence for hypothesized mechanisms.

In an effort to start addressing these issues, this study investigated reasoning questions in conjunction with peer collaboration and their affect on conceptual-knowledge learning. This study addressed potential mechanisms behind the efficacy of peer collaboration and reasoning questions through collection and analysis of process data in the form of collaborative student discourse. Using the perspective of collaboratively-regulated learning, analysis of these data provide evidence for the hypothesized mechanisms at work with questioning and peer collaboration.

The next section discusses the research on learning with hypermedia CBLEs, which was a contextual factor in this study.

#### Learning with Hypermedia CBLEs

In their reviews of learning with hypermedia, Dillon and Gabbard (1998) and Dillon and Jobst (2005) discuss the purported benefits of learning using hypermedia, including providing non-linear access to information, providing on-demand exploration of information, allowing self-paced instruction and increased engagement. However, both reviews reported scant evidence for these advantages over traditional learning environments in the research literature. Further, much research devoted to hypermedia learning in the past two decades has paid little attention to learning theory devoted to learning with hypermedia (Dillon & Gabbard, 1998; Dillon & Jobst, 2005). However,

multimedia learning, of which hypermedia learning is a type, does have a strong theoretical base, which is currently used in research on learning with hypermedia (Dillon & Jobst, 2005).

Multimedia describes information sources, usually computer environments, in which more than one representation of information is available, such as text, audio, video and still graphics (Mayer, 2005; Reed, 2006). Hypermedia is a type of multimedia in which information is connected through nodes called hyperlinks (Dillon & Jobst, 2005).

This section reviews theories of multimedia learning, which inform research on learning with hypermedia. I then address what reviews of research on learning with hypermedia have revealed in the past two decades, drawing heavily from the two seminal reviews on this topic by Dillon and Gabbard (1998) and Dillon and Jobst (2005).

*Theories of Learning with Multimedia and Hypermedia* Research on learning with multimedia and hypermedia has been guided by several theoretical perspectives. As Reed (2006) describes, these theories fall into one of two categories: multimodal theories investigating learning in traditional laboratory studies and instructional theories used in instructional contexts. The first group of theories includes Paivio's (1969) dual coding theory, Baddeley's (1974) working memory model, and Englekamp's (1998) multimodal theory. The second group of theories includes Sweller's (1994) cognitive load theory and Mayer's (2001) principles for learning with multimedia. These latter theories aim to explain how students learn, as well as how they best learn, using multiple modalities of information.

### *Multimodal Theories*

Paivio's (1969) dual coding theory was one of the first to distinguish between visual and verbal coding of information. He investigated these modalities in association with long-term memory (Reed, 2006). Research using this theory helped establish that pictures are more easily remembered than words, and that concrete words for which a mental image can be constructed are easier to remember than words that are abstract (Paivio, 1969). Baddeley and Hitch (1974) extended this line of research to include short-term memory and phonological coding in particular. Later instantiations of this theory have included a component (called an episodic buffer) in their theory that serves to explain how visual and verbal information are integrated. Engelkamp (1998) has taken these theories further and has included action into the theory. Engelkamp's research revealed that acting out phrases led to better recall of the phrase. He extended the theory by including the relation of the visual and verbal systems to the conceptual system and enactment. Essentially, to enact a phrase, one must understand it conceptually, which ensures it has been stored in memory.

The multimodal theories of multimedia learning are looking at learning with multiple modalities of information, yet their focus is just on the mechanisms by which memory of information is made. Sweller's (1994) and Mayer's (2001) instructional theories of multimedia learning have taken prior research in a new direction by focusing on theories that impact the design of multimedia to facilitate learning. Thus, the multimodal theories have become building blocks upon which recent research and theory has aimed to expand, in an effort to address multimedia learning in instructional contexts.

### *Instructional Theories*

Sweller's (1994) *cognitive load theory* is one that often guides research in instructional design of multimedia and hypermedia (Paas, Renkl, & Sweller, 2003). It assumes that students have a limited capacity in their working memory, which they use when solving problems or learning. When this working memory is overloaded, learners cannot construct the necessary schemas for actual learning and understanding. This is particularly true of novices who have not automated any of the procedural knowledge required for a given situation. As such, cognitive load theorists would argue that the manner in which information is presented is important and the expertise level of the learner should be taken into account (van Gog, Ericsson, Rikers, & Paas, 2005; Kozma, 2003). When learning with hypermedia, this theory can be considered particularly important, as hypermedia often offers students many choices and options, which can be overwhelming for many students.

Mayer (2001, 2005) provides a comprehensive theory of learning with multimedia, in that it synthesizes and expands the research and theories from Paivio, Baddeley and Sweller, in particular. His theory, based on years of research, has revealed a number of multimedia principles that impact students' learning when using multimedia (Mayer, 2001). The focus of Mayer's research has been how multimedia can best promote learning, particularly of mechanical and scientific processes. In particular, Mayer is interested in how different conditions within animation-based multimedia environments affect learning. The multimedia animations Mayer uses in his research have commonly been short, no more than several minutes in length, and serve to teach about simple mechanical process such as how lightning forms, how car brakes work, or how a tire pump works. More recent research has begun to explore botany and geology

multimedia environments (e.g., Mayer, Mautone & Prothero, 2002; Moreno & Mayer, 2005).

Mayer uses a combination of different theories to guide his research, the most prominent of which are information-processing theory, mental model theory, Witrock's (1974, 1989) generative theory, and Paivio's (1969) dual-coding theory (Mayer, 1997). Mayer's theory, based on these, is that "meaningful learning" (defined by good retention and transfer performance) happens when students select, organize, and integrate relevant information, and that these processes take place in two different systems, the visual and verbal systems.

The theory is based on inferences drawn from Mayer's research. For example, the results of the research done by Mayer under this framework demonstrate that students learn short, cause-and-effect descriptions about processes better if they are presented with a visual (e.g., an animation or pictures) as well as verbal (e.g., written text or narration) description than with verbal alone (as described in Mayer, 1997). Mayer called this result a *multimedia effect*.

Mayer's research also shows evidence of *contiguity effects*, whereby verbal information provided concurrently with visual information leads to greater scores on transfer tests than when the information is provided separately. The research also elucidated proof of a *split-attention effect* in multimedia learning, whereby learners perform better on transfer tests when the verbal form of information is presented as auditory narration versus on-screen text. This is presumably because the on-screen text overloads the visual processing that has incoming data from the visual information. By

presenting the information as auditory and visual, two systems are utilized and more information is taken in (Mayer, 1997).

Using the above-mentioned effects as guides, Mayer and his colleagues have developed a more comprehensive cognitive theory of multimedia learning with seven principles. There are three assumptions on which this theory is based: a) humans have two separate paths for processing visual and verbal representations (dual-channel); b) humans have a limited amount of space for processing at any given time (limited capacity); and c) humans learn best when they are actively constructing their knowledge by selecting, organizing, and integrating new knowledge in their working memory (active construction; Mayer, 2001). Mayer and colleagues' view is that a multimedia environment can foster the best development of a mental-model if the environment embodies Mayer's (2001) multimedia-design principles.

Mayer's (2001) seven self-stated principles for the design of multimedia environments are: the *multimedia principle*, the *spatial contiguity* principle, the *temporal contiguity* principle, the *coherence principle*, the *modality principle*, the *redundancy principle*, and the *personalization principle*. These principles of multimedia design for short, simple mechanical processes have been strongly supported by Mayer and colleagues' research, which has been guided and structured by Mayer's cognitive theory of multimedia learning and its attendant assumptions.

The spatial and temporal contiguity principles state that students learn better when they can access the visual and verbal data close together in space and time (Mayer, Moreno, Boire, & Vagge, 1999; Moreno & Mayer 1999). This allows their working memory to integrate the information more efficiently. The main assumption behind the

spatial contiguity and temporal contiguity principles is the dual-channel assumption. This assumption is based on the idea that we have two channels, visual and verbal, in our working memory that work together to make sense of incoming information.

The modality principle states that students learn better when one channel is not overloaded (i.e. visual) by having to read text *and* look at an animation – so auditory narration (with animation) is the preferable means of communication for the verbal processor (Mayer & Moreno, 1998; Moreno & Mayer, 1999; Moreno & Mayer, 2002b). The coherence principle states that extraneous material will overload working memory processors and therefore detract from learning (Moreno & Mayer 2000a; Mayer, Heiser, & Lonn, 2001). (This is in opposition to arousal theory, which states that extraneous music, sounds, and information might make a student more interested in a multimedia presentation, thereby resulting in further learning.) The redundancy principle states that providing redundant material results in less learning because it takes up valuable working memory space (Mayer, Heiser, & Lonn, 2001; Moreno & Mayer, 2002a; Moreno & Mayer, 2002b). The modality, coherence, and redundancy principles are mainly supported by the limited capacity assumption, though the dual-channel assumption is implicit.

The personalization principle was also derived from the limited capacity assumption. The thought is that personalization, in the form of informal rather than formal conversation, results in greater learning because it is more familiar and thus more easily processed by the verbal system (Moreno & Mayer 2002b). Other types of personalization, such as simple interaction, allow the learner to be more active and have

more control over the information they are receiving so that working memory capacity is regulated (Mayer & Chandler 2001).

Most recent research has revealed a potential new principle: the static media principle (Mayer, Hegarty, Mayer, & Campbell, 2005). Mayer et al. (2005) found that students presented with animation and narration did not outperform students presented with static images and text on transfer or retention tasks. In fact, across eight comparisons, the static group performed better than those in the dynamic task. This supports previous research by Hegarty, Quilici, Narayanan, Homquist, and Moreno (1999) who found that when learners are able to mentally animate mechanical systems from a static image, providing them with a visual animation did not lead to greater learning outcomes. Further research has shown that prompting learners to predict behavior of the system from a static diagram and providing them with a verbal description of the processes of the system facilitates learning (Hegarty, Kriz, & Cate, 2003). More research needs to be conducted to support these findings, but potential explanations for the effect may be that the static forms require more active processing by the learner (Mayer et al., 2005).

#### *Research on Learning with Hypermedia*

Dillon and Gabbard (1998) and Dillon and Jobst (2005) have written comprehensive reviews of the literature on learning with hypermedia. Dillon and Gabbard (1998) identified three focal areas of hypermedia research among the body of literature: learner comprehension, learner control, and individual differences in learning style. They found that the majority of studies that looked at differences in student comprehension between those who used paper-based information sources and those who

used hypermedia reported no significant differences in comprehension between the two groups of students. As such, recent research has turned to looking at specific aspects of hypermedia that may influence how students learn. The most researched topics have been the impact of the structure of the information within the environment, the role of advanced organizers, the impact of different levels of learner control, and the role of prior knowledge in learning with hypermedia. These studies have been conducted in wide variety of academic domains, including biology, ecology, physics, geology, history, psychology, foreign languages, business, and computer programming (Dillon & Gabbard, 1998; Dillon & Jobst, 2005).

Research on the structure of information within the hypermedia environment has primarily focused on comparing linear organization of information, which is similar to how a book is organized, to hierarchical structure. This has been a focus of hypermedia research because the linking of information is what makes hypermedia unique compared to other forms of multimedia. The research to date has revealed that structure does not appear to affect learning outcomes, but other variables, such as time on task, do seem to be affected (e.g., de Vries & de Jong, 1997). In particular, students tend to use less time with a hierarchical structure than with a linear structure (Dillon & Jobst, 2005).

Studies investigating the role of advanced organizers such as outlines, content lists, and conceptual maps, have revealed that advance organizers are helpful as compared to providing no advanced organizer (e.g., Shapiro 1999, 2000). However, the type of advanced organizer provided does not seem to make a difference (Dillon & Jobst, 2005).

Studies looking at the impact of increased learner control in hypermedia environments are based on the constructivist hypothesis that increased learner control will have a positive effect on learning because the learner can set his or her own path and pace (e.g., Jonassen & Reeves, 1996). Dillon and Gabbard's (1998) review found that increased learner control was only beneficial for high-ability learners; low-ability learners displayed difficulty using increased control features in a number of the reviewed studies. However, Dillon and Jobst (2005) report more current studies that found that increased control was better than limited control for most learners (e.g., Yeh & Lehman, 2001).

Individual differences such as prior knowledge have also been a focus of hypermedia research. This research has demonstrated mixed results, as prior knowledge often seems to interact with other variables such as the organizational structure of the information (Dillon & Josbt, 2005). For example, Calisir and Gurel (2003) and Potelle and Rouet (2003) found that low-prior knowledge students fare better with a hierarchical structure than with a more linear structure. Much of the research on this topic indicates that low-prior knowledge students tend to have greater difficulty learning with hypermedia than higher-prior knowledge peers, and this bears out in outcome measures (e.g., Lawless & Kulikowich, 1996; Recker & Pirolli, 1995). Dillon and Jobst (2005), drawing on discussions from studies they reviewed, suggest that students with low prior knowledge may not be able to effectively use links or nodes to learn because they do not already have a developed schema of the content to draw on.

#### *Hypermedia Learning and the Current Study*

The multimedia/hypermedia learning theories and research are pertinent to the current study for several reasons. First, hypermedia is the focal CBLE for this research.

The design of Encarta, the hypermedia environment students will use to learn about the circulatory system, follows many of the principles developed by Mayer and colleagues. For example, the static media principle is employed by depiction of static diagrams (e.g., heart and its chambers) along with written description of the processes (e.g., blood flow through the heart chambers). According to Mayer (2001) this design may facilitate mental animation of the process. However, the multimedia environments employed by Mayer and colleagues have largely consisted of displays of short, mechanical sequences. In contrast, the current study involved a much more complex environment consisting of multiple hyperlinks, text, and diagrams. Further, the tasks that learners engaged in in the majority of Mayer and colleagues' studies were fairly well-structured. In contrast, the task used in the current study was open-ended and ill-structured. The generalizability of the principles laid out by Mayer and colleagues to more complex environments and tasks is assumed, but there is as yet scarce evidence of this.

The nature of the well-constrained environments and tasks in Mayer and colleagues' research may provide one explanation for the focus purely on the hypothesized cognitive processes involved in learning with hypermedia, with no attention to other determinants of learning, such as metacognition and strategic behavior. There is little need for self-regulation with short, constrained tasks. While the theories of multimedia learning have certainly advanced our understanding of how best to design hypermedia and multimedia environments, the theories lack the explanatory power of the process of complex student learning with a hypermedia CBLE, even if that CBLE is designed based on well-researched principles such as those outlined by Mayer (2001). The cognitive architectural aspects (e.g., Baddeley's episodic buffer) do provide an

explanation for how stimulus information is transformed into knowledge stored in memory, but they do not explain the myriad other factors that impact students' ability to learn using hypermedia in educational contexts. While multiple modalities of information may make learning difficult, students also may be facing challenges in regulating their cognition, motivation, behavior, and context as they are learning (Azevedo, 2005; Winters, Greene, & Costich, 2008).

The research on hypermedia learning to date has scarcely addressed these factors. The lack of focus may be intentional, as Mayer (2001) even claims “[m]eaningful learning outcomes depend on the cognitive activity of the learner during learning *rather than on the learner’s behavioral activity during learning* [emphasis added] (p.1).” Other determinants of learning, such as planning and strategy use, are thus not included in the aforementioned theories, and in fact, are not viewed as a necessary aspect of learning with hypermedia under the reviewed theories. As such, the methodology used in most hypermedia research has focused predominantly on product data, usually in the form of scores on recall and transfer measures. If product data are collected, they have traditionally been in the form of navigation patterns rather than learning processes.

Self-regulated learning (SRL) provides a potentially more comprehensive framework for investigating how students are learning about complex topics while using hypermedia (Azevedo, 2005). In fact, SRL subsumes the multimedia theories, which help explain only part of the cognitive area of the SRL framework. The multimedia theories do not address other psychological areas of regulation, specifically, a students' motivation (e.g., self-efficacy judgments and interest) and behavior (e.g., monitoring of time use or help-seeking behavior), or the influence of the context (e.g., influence of

peers or teachers) on learning. Further, while the multimedia/hypermedia theories may address cognition, their focus is narrow, and they do not take into account the various phases of cognition, which include planning and activation (e.g., goal-setting), monitoring (e.g., judgments of learning), control (e.g., the strategy of note-taking), and reactions to the task afterwards.

The current study contributes to the theories of multimedia/hypermedia learning by empirically investigating the role of regulatory learning processes on conceptual-knowledge learning about complex science topics.

## CHAPTER 3: METHOD

### Participants

One-hundred and thirty three (N= 133) high-school students from secondary schools located in the mid-Atlantic region participated in the study. According to Cohen (1988), this is an appropriate sample size for a 2 X 2 factorial design with power equal to .80 and an alpha level of .05 when expecting a medium effect size for main effects. The students constituted a sample drawn from both public and private schools in the mid-Atlantic region.

The sample consisted of both male (n = 55) and female (n = 78) students, who had little previous exposure to the circulatory system in their high-school science classes prior to participating. However, most had some exposure during a prior middle-school health or science class according to demographic data, and as was evidenced in a pilot study conducted in the summer of 2006 with 10 representative students. This limited prior exposure was verified through consultation with the students' classroom teachers; it was further confirmed through analysis of pretest scores (M = 13.7%, SD = 10.6), which reflected students' prior knowledge about the circulatory system.

The sample was chosen from three schools in the mid-Atlantic region. Two of the schools were private high schools, from which 39% (n = 53) of the sample was drawn. The other school was a larger public high school from which 61% (n = 80) of the sample was drawn. Data on ethnic diversity for all schools were comparable, with approximately 30% of the student body comprising minority students. The majority of participants from

both types of schools were juniors in high school (66%), with seniors comprising the next largest group (31%), and sophomores the smallest group (2%). No freshman participated in the study. The mean age of the students in the sample was 17 years old ( $SD = .740$ ). The majority of students (80%) expected a grade of B or higher as their final course grade.

### Research Design

The research design was a mixed-method design, including both quantitative and qualitative analyses. The quantitative portion of the design was a 2 (Learning condition: Peer learning vs. Individual learning) X 2 (Questioning condition: Question vs. No question) factorial design. The dependent variable was a measure of conceptual knowledge administered prior to and following the intervention (see Appendix B). The independent variables were the learning and questioning conditions, with two levels in each. The four experimental conditions were the No Question and Peer learning (NQP;  $n = 41$ ), the Question and Peer learning (QP;  $n = 36$ ), the No question and Individual learning (NQI;  $n = 28$ ), and the Question and Individual learning (QI;  $n = 28$ ). The qualitative portion of the design consisted of discourse analysis from audiotapes of a randomly assigned subset of dyads working together ( $n = 27$  dyads) in the peer-collaboration condition during the experimental task.

### Materials

#### *Conceptual-Knowledge Measure*

To assess students' conceptual knowledge of the circulatory system, they were administered identical paper-based pretests and posttests. The test was adapted from that used by Azevedo and colleagues (Azevedo, Cromley, & Seibert, 2004; Azevedo,

Cromley, Winters, Moos, & Greene, 2005) to gauge a students' conceptual knowledge of the circulatory system in the form of mental models. The adaptation includes elements of a structure-behavior-function (SBF) analytic approach for assessing conceptual knowledge of complex systems (Hmelo et al., 2000; Hmelo-Silver & Pfeffer, 2004).

The conceptual-knowledge measure included six constructed-response questions that were developed based on key concepts about the circulatory system (Chi et al., 1994; Towle, 2000). Those include: an understanding of the functions (purposes) of the circulatory system; structures (parts), which are the blood, blood vessels and heart; and, an understanding of structural interactions during the path of blood flow through the heart, lungs and body (see Appendix A). These structures and functions comprise key components of a mental model of the human circulatory system. As such, questions 1 and 2 ask students to list all the functions and parts, respectively, of the circulatory system. Questions 3 and 4 ask students to list the parts of the blood and blood vessels, respectively, then fill in coordinating functions with each part listed. Questions 5 and 6 ask the student to label a diagram of the heart and then describe the path of blood flow through it and the rest of the body (see Appendix A). Reliability estimates of the pretest (Cronbach's  $\alpha = 0.69$ ) and posttest scores (Cronbach  $\alpha = 0.76$ ) were acceptable.

#### *Development of the Measure*

The conceptual-knowledge measure was developed in several iterations. In the summer of 2006, I conducted a pre-pilot study with six participants to determine the best format for measuring of students' conceptual knowledge about the circulatory system. The preliminary formats included a selected-response, multiple-choice test; a free-recall, open-ended essay; and a structured interview. The purpose of this pre-pilot was to

determine if results similar to the free-recall essay that Azevedo and colleagues (Azevedo, Cromley et al., 2004; Azevedo et al., 2005) have used to capture conceptual knowledge could be obtained with a selected-response, multiple-choice measure of conceptual knowledge. The interview was used to confirm that the free-recall essay was assessing what students knew about the circulatory system well, and responses were scored using the same method as the free-recall essay (see Azevedo, Cromley et al., 2004).

Of interest was whether the selected-response measure could be used instead of the free-recall essay to assess conceptual understanding. The selected-response measure required less writing, and of concern was the writing ability of the expected participants, which were 9th and 10<sup>th</sup> grade students. From an evidence-centered assessment design perspective (ECD, Mislevy, Almond, & Lucas, 2004), writing ability would constitute an additional skill that is required for the student to complete the assessment and could compromise the validity of the results of the scores if not taken into account (Mislevy & Riconscente, 2006). If students are not able to write well, the essay would not necessarily be measuring conceptual knowledge accurately.

The results of this pre-pilot were informative in several ways. First, for each student, the conceptual knowledge score varied with each of these forms of assessment, indicating they could not all be considered equally valid measures of conceptual knowledge. Across the six participants, the selected-response measure provided the least information about what students knew, and it did not discriminate well between students, particularly as compared to the other two measures of conceptual knowledge. Second, it

was also apparent from the essay responses with the pre-pilot that all the students were, in fact, comfortable writing a constructed response.

Third, the interview, which provided more specific prompts (e.g., *what are the parts of the circulatory system*) than the free-recall essay's more open-ended prompt (i.e., *write down all that you know about the circulatory system*), provided the most information from each student about their conceptual understanding of the circulatory system. In other words, the high-school students in the pre-pilot sample provided more information about what they knew when given a specific prompt than when they were given an open-ended prompt.

As such, the conceptual-knowledge measure used in a subsequent pilot consisted of six constructed-response items to specific prompts about the parts and functions of the circulatory system. Specifically, questions 1 through 5 asked students to write down the functions of the circulatory system, the parts of the circulatory system, the parts and functions of the heart, the parts and functions of the blood, and the parts and functions of the blood vessels. Question 6 asked students to use a diagram to describe the path of blood flow through the heart and body. These items were developed to tap similar components identified as comprising conceptual understanding by the free recall essay used by Azevedo and colleagues (Azevedo Cromley, et al., 2004; 2005). As such, the total score on the measure is considered a measure of the students' conceptual understanding. The items could be parsed, with items 1-5 assessing declarative knowledge about parts and functions individually and item 6 assessing knowledge about how these parts and functions interact in the process of blood flow.

### *Pilot Test of Measure*

Ten pilot students were administered the new pretest and posttest, with a 30-minute learning session about the circulatory system in between. The measure discriminated well between students, but six out of the 10 students took longer than 20 minutes to complete the posttest (range: 9 – 30 minutes). For practical data collection purposes, a time limit of 20 minutes is needed for the pretest and posttest. As such, to decrease the time needed for students to complete the posttest, three of the items were altered to minimize the amount of writing needed. First, question 3, which asked students to write down the parts and function of the heart was replaced with a heart-labeling item, which requires knowledge about the parts of the heart. Questions 4 and 5 were changed from a short-answer format to filling in a table with single words or short phrases. The new format elicits the same information, but minimizes the need for writing sentences (see Appendix A).

A time limit of 20 minutes was given to students for both the pretest and posttest. The rationale for this particular time limit was: a) to force students to be succinct, without penalizing students who may not write as quickly as others (based on the pilot study, most students should be able to complete it within 20 minutes); and b) to fit a school's class scheduling, within which this study was conducted.

### *Hypermedia Learning Environment*

The students used Microsoft's Encarta Encyclopedia (2004), a multimedia encyclopedia, to learn about the human circulatory system. It has been used in a number of studies conducted by Azevedo and colleagues, with students from 7<sup>th</sup> grade through college level (e.g., Azevedo, Cromley, et al., 2004; Azevedo et al., 2005; Moos &

Azevedo, 2006). The learning environment contains many articles, including three main articles about the circulatory system: *Circulatory System*, *Heart*, and *Blood*, that students were guided to use. These articles contain 16,900 words, 35 illustrations, 107 hyperlinks, 18 sections, and 3 main articles.

### *Reasoning Question*

At the start of the task, students in the questioning condition were provided with a paper sheet on which a reasoning question was written. Particular types of questions are considered reasoning questions when answering them requires explanatory reasoning, in the form of logical syllogisms or causal chains, as has been evidenced with tutor-student interactions (Graesser et al., 1996; Graesser & Person, 1994). This assumption was tested and validated by analysis of the written responses provided by the students. Below each question, there was space for the students to write their answers (see Appendix B).

The question reads:

Let's pretend you were a blood cell, located for the moment in a healthy person's big toe – inside the right or left foot (in this case it makes no difference). It would take you longer to travel from the big toe up to the left side of the heart than it would for you to get from the left side of the heart back down to the big toe.

Why? Explain as many reasons possible for why this is true.

Under the Graesser and colleagues' coding scheme for types of reasoning questions (Graesser et al., 1992; Graesser & Person, 1994) this type of question is considered an *antecedent question*, because it asks students to consider the way in which the blood cell must move around the body to return to the same spot. It should elicit causal chains, as a

form of reasoning, because it requires students to progress through a logical series of causal events in the process to answer the question correctly.

The answer to the question could not be directly obtained from the text, so students had to make inferences from what they have learned to arrive at the answer.

### *Pilot Test*

I conducted a pilot study with 5 individual students using a similar, but less complex reasoning question that asked students to trace the path of blood flow from the toe through the heart and back again. One concern was that students might answer the question by writing verbatim from a particular paragraph or section in the environment, rather than constructing an explanatory answer on their own. Analysis of student responses revealed that students did not use one particular paragraph, but rather drew from several sources, putting answers into their own words rather than quoting sections directly from the text. One participant did not address the question, having misunderstood the directions to answer them within the 30 minutes. She later indicated that she had planned to work on it after the 30 minutes were over. For this study, I ensured that the participants used the prompt as such by stating clearly in the instructions that the question was to be answered within the 30-minute task and by providing time reminders.

### Procedure

Parental consent forms were handed out in advance to students and collected prior to the start of students' participation in the study. Students whose parents consented were also given an assent form to sign prior to their participation in the study. Student participants were then randomly assigned to either the Question (Q) or No Question (NQ) condition, as well as to working with a peer (P) or working individually (I). Those

students randomly assigned to work with a peer were assigned to work with another student of the same gender to mitigate any possible influence on student interactions related to gender (e.g., Webb, 1984). The students worked at laptop computers in a school classroom or on desktops in the school's computer lab, during the students' regularly-scheduled class time.

All students were given 20 minutes to complete the paper-based pretest (see Appendix A). Students were informed that they would also be completing a posttest, but no more information about the contents of the posttest was provided (e.g. that it is identical to the pretest). Students were given a short tour of the learning environment to demonstrate the articles available, how to navigate through articles, and how to use the search feature.

Following the tour, students were provided with directions for their task. There were four different conditions, with corresponding sets of instructions, as described in Table 1.

The classroom teacher and I circulated at the beginning of the session to ensure that students understood the instructions provided. Students in all conditions were told that they could take notes during the task, but that they would not be able to use them (or the question and answer) during the posttest. The posttest was a measure of how much the students learned while participating in the task, and the use of aids such as notes or answers to the question while taking the test would undermine the assessment of their learning. Students then proceeded to the learning environment where they spent 30 minutes learning about the circulatory system. Students were given paper and pens to

Table 1  
Matrix of Instructions, by Condition

	Question	No Question
Individual Learner	<p>You will be using Encarta to learn about the circulatory system today. Your task is to learn all you can about the circulatory system in 30 minutes. <i>Make sure you learn about the different parts and their purpose, how they work both individually and together, and how they support the human body.</i> You will have a test afterwards to assess what you have learned.</p> <p><i>You are being given a question designed to help you learn today. They will be available to you throughout the learning task, and there is space below each question for you to write your answer. This question is to be answered during the 30-minutes you use Encarta.</i></p>	<p>You will be using Encarta to learn about the circulatory system today. <b>Your task is to learn all you can about the circulatory system in 30 minutes.</b> <i>Make sure you learn about the different parts and their purpose, how they work both individually and together, and how they support the human body.</i> You will have a test afterwards to assess what you have learned.</p>
Peer Learners	<p>You and your partner will be using Encarta to learn about the circulatory system today. Your joint task is to learn all you can about the circulatory system in 30 minutes. <i>Make sure you learn about the different parts and their purpose, how they work both individually and together, and how they support the human body.</i> You will have a test afterwards to assess what you have learned.</p> <p>You must work together to help each other learn the material.</p> <p><i>As a pair, you are also being given a question designed to help you learn today. They will be available to you throughout the learning task, and there is space below each question for you to write your answer. This question is to be answered during the 30-minutes you use Encarta.</i></p>	<p>You and your partner will be using Encarta to learn about the circulatory system today. <b>Your joint task</b> is to learn all you can about the circulatory system in 30 minutes. <i>Make sure you learn about the different parts and their purpose, how they work both individually and together, and how they support the human body.</i> You will have a test afterwards to assess what you have learned.</p> <p>You must work together to help each other learn the material.</p>

take notes if they chose to do so, as well as for answering the reasoning question, for the students in that condition.

Students in the Question condition were told that they could answer the question whenever they wished during the 30 minutes. The reasoning question was given to students in this condition at the beginning of the task, rather than in the middle or at the end. This placement decision was made because research has demonstrated that adjunct pre-questions are better at facilitating positive academic outcomes than adjunct post-questions when students are given a time limit (Hamaker, 1986), as in this study. The question was available for the duration of the learning task (i.e., 30-minutes of using Encarta to learn about the circulatory system). A sample of 27 peer-learning dyads (out of 39 total), were audiotaped while they collaborated. This constituted the collection of process data.

When there were 10 minutes left in the task, I alerted students to the pending time limit. At the completion of the learning task, all of the students were given 20 minutes to individually complete the posttest, which was identical to the pretest.

After participating, each class was debriefed as a whole. They were told that conceptual-knowledge learning was being compared between the different experimental groups to determine the effects of peer-collaboration and questioning. As well, they were told that analysis of recorded peer discourse would be analyzed to find relations with learning outcomes. Table 2 presents the experimental procedure for the questioning and no questioning conditions and for the peer and individual conditions.

Table 2

*Experimental Procedure for Questioning and No Questioning Conditions, for Peer and Individual Learners*

<b>Experimental Procedure</b>					
<b>No Questioning Conditions : Peer and Individual</b>	<b>Pretest</b>	<b>Tour of Encyclopedia</b>	<b>Instructions</b>	<b>Learning task</b>	<b>Posttest</b>
<b>Questioning Conditions : Peer and Individual</b>	<b>Pretest</b>	<b>Tour of Encyclopedia</b>	<b>Instructions + Reasoning Question</b>	<b>Learning task + Reasoning Question</b>	<b>Posttest</b>
<b>Time (min)</b>	<b>Max.20</b>	<b>2</b>	<b>1</b>	<b>30</b>	<b>Max. 20</b>
<b>Product Data Collection</b>	*			*	*
<b>Process Data Collection</b>				←*****→ For Dyads only	

*School Environments*

Despite the identical procedures carried out at each school, there were differences between the schools and the circumstances surrounding the data collection at each. The private schools were similar to each other, with small class sizes (approximately 12-15

students per class), computers or laptops in the classroom, and students had an evident comfort with the teacher and the instructional space. Data collection in one of the schools was conducting in one block-scheduled (90 min.) class period, whereas for the other school, the students took the pretest two days prior to engaging in the task and posttest. Other than this difference, the climate of the classes felt similar between the two private schools. During the data collection in the private schools, the students' teachers did not directly interact with the task or the students after helping get them set up and started, and they usually worked in other areas of the classroom or left for extended periods of time. I was considered the guest teacher for that period by the students and their teacher.

The public-school classes were nearly twice as large as the private-school classes (approximately 24-26 students). The public school students took the pretest one day prior to engaging in the task and posttest. Unlike the private-school students, the task and posttest were conducted in the school's computer lab. This was a considerably more crowded space than the private-school students had when working on the task in their classrooms. However, the students appeared to be familiar and comfortable working in this space. The larger class size meant a greater noise level as students worked. Similar to the private school data collection, the teachers in the public school treated me as the guest teacher and often left the computer lab for extended periods of time. As a consequence, I spent some of my time doing a classroom management and expended effort getting students set up with the task— a situation I did not encounter to such a degree in the private-school settings.

## Coding and Scoring

This section describes the coding and scoring of the students' conceptual knowledge (product data), the segmentation of the students' verbalizations, the coding scheme used to analyze the students' discourse (process data), the coding and scoring of students in the questioning condition's answers to the question, and procedures for ascertaining interrater agreement. Table 3 presents a summary table of the coding and scoring procedures.

### *Conceptual-Knowledge Measure*

The product data analyses focused on the students' conceptual knowledge at pretest and posttest. The coding procedure was adapted from Azevedo and colleagues coding of mental models (Azevedo & Cromley, 2004; Azevedo, Cromley, et al., 2004; Azevedo et al., 2005) and the SBF analysis framework (Hmelo et al., 2000; Hmelo-Silver & Pfeffer, 2004). The measure consists of six constructed-response questions. They were designed to assess students' knowledge about the parts and functions of the circulatory system, as well as how they work together. This information constitutes the knowledge necessary for comprehensive conceptual-knowledge of the circulatory system. The first five questions asked students to recall the functions and structures of the circulatory system, as well as describe the parts and functions of the major structures, which are the heart, blood, and blood vessels. Question 6 asked students to describe the path of blood flow through the heart and body using an unlabeled diagram. Students' answers to the six constructed response questions were scored based on the number of key structures and functions provided in their answers (see Appendix C for scoring rubric).

Table 3

*Summary of Coding and Scoring Procedures*

<i>Data Source</i>	<i>Description</i>	<i>Coding/Scoring</i>
<i>Conceptual-Knowledge Measure (Appendix A)</i>	<p>Pretest and Posttest: 6 constructed response questions</p> <p><b>Questions 1-5:</b> Structures and Functions of Circulatory System (Heart, Blood, Blood Vessels)</p> <p><b>Question 6:</b> Path of blood flow through body, heart and lungs (e.g., mental model of circulatory system, Azevedo, Cromley, et al., 2004)</p>	<p>1. Score questions based on rubric in <b>Appendix C</b></p> <p>Range of scores = 0 – 59</p>
<i>Peer Discourse</i>	Audio tapes of peers learning together	<p>1. Transcribe audiotapes;</p> <p>2. Segment transcriptions based on utterances (self-contained thought, idea or statement);</p> <p>3. Code utterances based on collaborative self-regulatory coding scheme in <b>Appendix D</b></p>
<i>Question Responses (Appendix B)</i>	Written responses to reasoning question	<p>1. Score responses based on rubric in <b>Appendix E</b></p> <p>Range of scores = 1 – 8</p>

The key structures and functions, and their relative importance, were derived from the coding scheme of Azevedo and colleagues (2005), which was used in both the development of the measure and the scoring tool. The total points possible for the conceptual-knowledge measure were 60 points. The pilot study confirmed the adequacy

of this coding scheme for assessing students' knowledge of the parts and functions of the circulatory system.

### *Peers' Regulatory Discourse*

A sample of the student pairs ( $n = 27$  pairs) were audiotaped to capture their discourse and behavior as they worked on the task. These discourses were then transcribed from the audiotapes. The transcriptions were analyzed using a coding scheme developed by Winters and Azevedo (2005) for analysis of collaboratively self-regulated learning. The coding scheme is derived from Azevedo, Cromley, et al. (2004) and Azevedo et al.'s (2005) coding scheme for self-regulated learning within the particular context of learning with hypermedia. This scheme has been adapted to investigate students' regulation of their learning in a collaborative CBLE- based ecology unit (Azevedo, Winters, & Moos, 2004), as well as the collaborative discourse of students working in pairs with a genetics computer simulation in the Winters and Azevedo (2005) study.

The current scheme is a combination of the Winters and Azevedo (2005) scheme and the coding scheme used with hypermedia environments by Azevedo, Cromley, et al. (2004) and Azevedo et al. (2005), placed within a social-cognitive theoretical context in conjunction with the framework described by Pintrich (2000) for self-regulated learning. The codes include aspects of planning (i.e., forethought), control (i.e., performance), monitoring (i.e., reflection) and motivation. The codes are intended to identify particular behaviors—both facilitative (i.e., making an inference) as well as inhibitory (i.e., off-task talk)—that students engage in as they regulate their learning. However, the coding scheme itself does not make a distinction in the quality of the behavior identified (i.e., “good” as

compared to “bad” processes, or facilitative as compared to inhibitory processes). In this study, the subsequent qualitative analyses provided that role.

In the planning category, there are five codes. Prior knowledge activation (PKA) is evidenced when students articulate something they have learned previous to the task. Recycling a goal (RG) is coded when students repeat a learning goal that they have previously stated. Stating a sub-goal (SG) is coded when students state a learning goal other than the primary task goal or reasoning question. Teacher-set goals (TSG) are coded when learners verbalize task-set goals or the question on the worksheet. Time and effort planning (TEP) is coded when students articulate some sense of planning relative to the time they have left.

The performance category is the one with the largest number of codes ( $n = 14$ ). Codes in this category identify strategies students use to learn collaboratively using a hypermedia environment. They include coordinating information sources (COIS), which occurs when students use related multiple representations (e.g., text and a diagram) together. Making an inference (INF) is another strategy. It is defined as drawing a conclusion from two or more pieces of information from within the environment. Knowledge elaboration (KE), which is a similar strategy, involves elaborating on or explaining what was just read or seen using implicit prior knowledge. Memorizing (MEM) is coded when a student tries to memorize part of the text or diagrams, such as using a mnemonic device. Partner questioning for procedure (PQP) is coded when learners ask their partners a question pertaining to the procedure of the task. Reading out loud (RL) is coded when one partner reads the text out loud. Reading notes (RN) occurs when one student reads over his or her notes. Seeks consensus (SC) occurs when students

seek their partner's agreement on a conceptual or procedural decision. Searching (SEARCH) is coded when students articulate that they are searching the learning environment. Summarizing (SUM) is coded when a student puts something they have just read or looked at in their own words. Taking notes (TN) is coded when a student writes down information they are learning.

For monitoring activities, there are 15 codes. These codes identify various monitoring processes that occur during the learning session. Content evaluation positive or negative (CE + / -) is apparent when students state that content they have just viewed is relevant (+) or not (-) to their learning goal. Expectation of adequacy of content (EAC) occurs when students state an expectation that particular content will be relevant or not to their goal. Feeling of knowing positive or negative (FOK +/-) occurs when students articulate that something they have just read or seen is familiar (+) or not familiar (-). Judgment of learning positive or negative (JOL+/-) is coded when students indicate that they understand (+) or do not understand (-) what they have just read or seen. Monitoring progress toward goals (MPTG) occurs when students articulate an assessment of whether they have met a goal or not. Partner questioning for understanding (PQU) is coded when learners ask their partners a conceptual question, indicating they do not understand. Seeking affirmation (SA) is coded when students seek reassurance or agreement from their partners, usually after questioning the partners about a conceptual or procedural aspect of the task. Self-correction (SEC) is coded when a student states that he or she made a mistake or misjudgment. Lastly, time monitoring (TM) occurs when students refer to the number of minutes they have remaining in the learning session.

Finally, I include a fourth category related to motivation, which contains five codes. Based on the results of Winters and Azevedo (2005) and Azevedo, Winters, et al. (2004) studies, students made very few verbalization related to motivation. However, particular codes, such as off-task behavior, are critical to gathering evidence for reasons behind successful (or unsuccessful) collaborative learning. As such, these codes are included in the analysis of data for the proposed study. The codes in this category include off-task (OT) behavior, which is coded when learners exhibit behavior that is not within the learning task. Positive feedback (PF+) occurs when learners give encouragement to their partner in response to an idea, choice or question. Negative feedback (NF) is coded when learners give discouragement to their partner in response to an idea, choice or question. Positive interest (INT+) is coded when students express interest in what they are reading or learning, and disinterest (INT-) is coded when they articulate a lack of interest in what they are reading or learning.

The analysis involved segmenting utterances in the transcriptions and applying regulatory codes to those utterances that provide evidence for regulatory processes (see Appendix D). Figure 1 is an example of part of a coded transcript from the Winters and Azevedo (2005) study. Transcripts for the present study were similarly coded. Once a transcript was coded and then re-coded by another rater, the codes were tallied for each participant as well as for each dyad, constituting two separate units of analysis.

---

Turn Transcript

---

- 1     **Student H:** / I remember like last time we went up to generation 5, it took  
a long time /[FOK+].../Ok, so it's 17, 4, and 10/  
[SUM]...../That is kinda [unintelligible]/ [UC].
- 2     **Student L:** /Uh.....For the last one we say yes/ [PQP], /right/[SA]?
- 3     **Student H:** /Uh, what?[NC] /Yeah/[PF].....
- 4     **Student L:** /Uh, does that count/ [PQP]?
- 5     **Student H:** /[unintelligible]/ [UC]
- 6     **Student L:** /Is that answer yes/ [PQP]?.....
- 7     **Student H:** /All right, what is the next question/ [MPTG]?
- 8     **Student L:** /*If you set the mutation rate to zero, will dragons still grow  
legs in order to reach the land and get out of the water/ [TSG]?  
/I say no/[INF]-*

Note: Note: slashes (/) indicate beginning and end of an utterance; italicized words indicate reading aloud; non-italicized words are other utterances; bracketed words are SRL codes for the utterance immediately preceding the brackets.

[FOK+] = Feeling of knowing +

[SUM] = Summary

[UC] = Uncodable

[PQP] = Partner questioning for procedure

[SA] = Seek affirmation from partner

[PF] = Positive feedback

[MPTG] = Monitoring progress towards goals

[TSG] = Teacher-set goals

[INF] = Inference

*Figure 1.* Example of coded transcript from Winters and Azevedo (2005).

### *Question Responses*

Students in the questioning condition were asked in the written instructions to provide written answers to their question. To verify the quality of answers that students provide, the answers were scored on a scale from 0 – 8 based on the amount of detail and information provided in students' answers, which reflects their reasoning. A score of 0 indicated no answer given, while a score of 8 indicates a complete answer. For example, if a student provided an answer to the question in which she stated that the gravity slows the blood down going up, she received a score of 2. If, however, she explained that traveling from the left ventricle, the blood would have to make a trip to the lungs and then back to the heart before going back down to the toe, and that the pressure on the way

down is greater than on the way up, she received a score of 6 (see Appendix E for the full coding scheme).

#### *Interrater Agreement*

I was the primary coder, coding 100% of the conceptual-knowledge measures, the transcripts, and the question responses. Interrater agreement was established by training a second reader to code the conceptual-knowledge pretests and posttests and question responses. This second reader recoded 30% of the tests ( $n = 80$ ) and question responses ( $n = 20$ ) to ensure coder agreement. There was agreement on 72 of 80 test scores, yielding an interrater agreement of 91% for the pretests and posttests. There was agreement on 17 of 19 question response scores, yielding interrater agreement of 92% for the question responses. Similarly, a second coder who has six years of experience coding transcripts using Azevedo and colleagues' coding scheme was trained to use the adapted co-regulation coding scheme. He recoded 30% ( $n = 8$ ) of the participants' transcripts. There was agreement on 1037 of 1070 coded utterances, yielding interrater agreement of 96%. The second raters were blind to conditions. Inconsistencies were resolved through discussion between the raters and me.

## CHAPTER 4: RESULTS

### Descriptive Statistics

Before conducting parametric statistical analyses on the collected data, skewness and kurtosis values were analyzed for the pretest, posttest, and question response scores. To determine the range of values that would indicate normality, the standard error of skewness (*ses*) and kurtosis (*sek*), which are dependent on sample size, provided the appropriate interval within which the values should fall. Specifically, the absolute value of the skewness and kurtosis values needed to be below two times the standard error to indicate normality (Tabachnick & Fidell, 1996). Table 4 shows the skewness, kurtosis, standard errors, means, and standard deviation values for the pretest, posttest, gain, and question response scores overall, and for each subgroup that was used in subsequent analyses.

Analysis of the skewness and kurtosis values with respect to their standard errors revealed that the pretest scores were non-normally distributed for the total sample, as well as for the majority of the subgroups. As a result, the pretest scores were transformed for use in the subsequent analyses of variance using a square root transformation, which is appropriate for distributions that are moderately skewed (Tabachnik & Fidell, 1996). Question response scores had a kurtosis value for the total sample that indicated it was not normally distributed. However, for each subgroup, the scores were normally distributed. Because subsequent analyses were conducted on these subgroups only, the question response scores were not transformed.

The verbal protocol data that were collected are frequency data, and, by nature are usually not normally distributed. To avoid the loss of information that occurs when data are transformed, non-parametric analyses were used with these data. Of further interest in

Table 4  
*Means, Standard Deviations, Skewness, and Kurtosis Values of Study Measures*

	Pretest (%)	Posttest (%)	Gain (%)	Quest. Response Score
Total Sample <sup>1</sup>				
<i>M</i>	13.49	32.73	19.24	2.67
<i>SD</i>	10.40	18.68	14.74	2.05
Skewness	1.02*	0.44	0.44	0.07
Kurtosis	1.03*	-0.08	-0.52	-1.30*
NQP <sup>2</sup>				
<i>M</i>	13.93	36.40	22.47	--
<i>SD</i>	10.80	20.17	14.84	--
Skewness	1.34*	0.61	0.31	--
Kurtosis	2.80*	0.29	-0.38	--
QP <sup>4</sup>				
<i>M</i>	12.52	32.30	19.77	2.78
<i>SD</i>	8.81	17.56	13.68	2.10
Skewness	0.66	0.10	0.35	-0.04
Kurtosis	0.31	-0.79	-0.41	-1.34
NQI <sup>5</sup>				
<i>M</i>	14.59	30.84	16.25	--
<i>SD</i>	11.70	18.33	13.88	--
Skewness	1.10*	0.62	0.34	--
Kurtosis	0.53	0.63	-0.99	--
QI <sup>4</sup>				
<i>M</i>	12.98	29.78	16.80	2.54
<i>SD</i>	10.74	18.30	16.38	2.00
Skewness	0.62	0.21	0.87	0.87
Kurtosis	-0.77	-1.13	-0.11	-1.16
By School Type				
Public <sup>5</sup>				
<i>M</i>	14.07	25.52	11.45	1.90
<i>SD</i>	10.47	16.61	10.45	1.91
Skewness	0.85	0.90	0.63	0.63
Kurtosis	0.12	1.06	0.20	-0.78
Private <sup>6</sup>				
<i>M</i>	12.62	43.60	30.99	3.88
<i>SD</i>	10.33	16.32	12.30	1.64
Skewness	1.32*	0.16	0.04	-0.59
Kurtosis	2.96*	0.66	-0.49	-0.26

\*Indicates values that exceed those of normally distributed data for that sample size

<sup>1</sup> N = 133; *Ses* = .21; *Sek* = .43; <sup>2</sup> NQP = No Question, working with a peer; n = 41; *Ses* = .38; *Sek* = .77

<sup>3</sup> QP = Question, working with a peer; n = 36; *Ses* = .41; *Sek* = .82

<sup>4</sup> NQI = No Question, working individually; n = 28; *Ses* = .46; *Sek* = .93; <sup>5</sup> QI = Question, working individually; n = 28; *Ses* = .46; *Sek* = .93; <sup>6</sup> n = 80; *Ses* = .27; *Sek* = .55; <sup>7</sup> n = 53; *Ses* = .34; *Sek* = .67

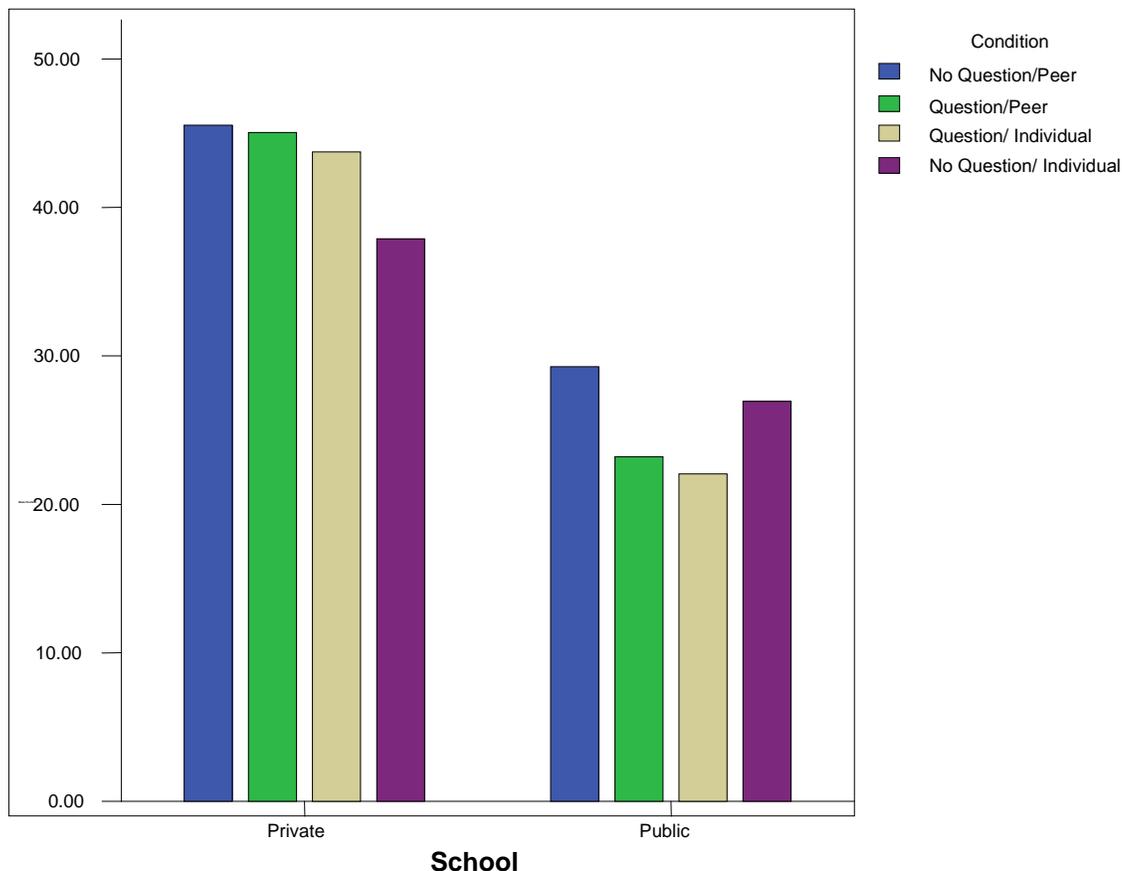
this study was the proportion of use of regulatory processes. Non-parametric analyses are appropriate for analyzing proportion.

## Research Question Results

### *Research Question 1*

Does working with a peer facilitate students' conceptual-knowledge learning about a complex science topic to a greater degree than learning alone? On average, students across all conditions made significant conceptual-knowledge gains from pretest ( $M = 13.49$ ,  $SD = 10.40$ ) to posttest ( $M = 32.73$ ,  $SD = 18.68$ ),  $t(132) = -19.08$ ,  $p < .05$ . However, visual inspection of the data revealed an unexpected potential effect of school type on pretest and posttest scores. T-tests between the private and public schools on pretest and on posttest scores revealed no significant differences on the pretest scores between the two school type groups,  $t(131) = -0.98$ ,  $p > .05$ , but significant differences on the posttest scores,  $t(131) = 6.19$ ,  $p < .05$ , between the two groups. Figure 2 shows this difference graphically, by experimental condition.

The mean difference in gain from pretest to posttest between the two school groups was 18.08 percentage points, with private school students evidencing the greater mean value ( $43.60$ ,  $SD = 16.32$ ) than the public school students ( $25.52$ ,  $SD = 16.61$ ) at posttest. A t-test on the gains from pretest to posttest for each school type revealed significant differences for both the private schools,  $t(52) = -19.27$ ,  $p < .05$ , and for the public school  $t(79) = -12.80$ ,  $p < .05$ .



*Figure 2.* Mean posttest scores by experimental condition and school type.

Despite school type differences, the variability within each school type on posttest was nearly equivalent to the variability between them. As such, school type was not added as an effect in the subsequent analysis of covariance. A 2 (Peer condition) X 2 (Question condition) analysis of covariance, with pretest as the covariate and posttest as the dependent variable, revealed a non-significant main effect for the Peer condition,  $F(1, 128) = 2.39, p = .12$ , indicating that those learning with a peer did not have significantly higher adjusted posttest scores than those learning individually at the  $p < .05$  level of significance. There were no significant interaction effects. See Table 5 for a summary of the results of the ANCOVA.

Table 5

*Analysis of Covariance for Peer and Question Conditions*

Source	df	F	$\eta$	p
Pretest <sup>1</sup>	1	65.03	0.34	0.00
Peer	1	2.39	0.02	0.12
Question	1	0.14	0.00	0.71
Peer x Question	1	0.37	0.00	0.54
Error	128	(74.79)		

<sup>1</sup>Covariate, transformed via square-root transformation

Note: Value enclosed in parentheses represents mean square error

*Research Question 2*

Do peers engage in collaborative self-regulatory processes as they learn together?

Is there evidence that these processes are related to learning outcomes? The results for the first part of this research question indicated that students do engage in collaborative self-regulatory processes as they learn together. Table 6 shows the mean and median percentages of codable utterances for each of the regulatory categories over all audiotaped students. Median percentage is reported as well as mean percentage because the data were not normally distributed. As Table 6 indicates, approximately 90% of the utterances for all audiotaped students could be classified under the collaborative self-regulatory coding scheme.

The highest percentage of codable utterances was in the category of performance (46%), which consisted of learning strategies, and the lowest percentage of utterances was the forethought category (7.2%). Approximately 9% of students' utterances were uncodable. These utterances consisted of unintelligible utterances, comments about the

hypermedia environment (e.g., *look, that changes to pink up there*), and other comments that did not provide clear evidence of a regulatory behavior (e.g., *yeah...[6 second pause]...ooh, pulmonary*).

Table 6

*Mean, Standard Deviation, and Median Percentages of Collaborative Self-Regulatory Learning Processes Overall Audiotaped Pairs and by School Type*

	Forethought	Performance	Monitoring	Motivation	Uncodable
Overall					
Mean	8.4	44.6	15.9	22.0	9.5
(SD)	(6.8)	(15.1)	(6.9)	(17.0)	(6.5)
Median	7.2	46.1	16.7	16.7	9.3
Private					
Mean	6.6	50.4	15.5	18.9	8.5
(SD)	(4.5)	(13.9)	(7.7)	(15.7)	(6.1)
Median	6.5	50.0	16.3	15.8	9.0
Public					
Mean	10.6	37.1	16.3	25.9	10.5
(SD)	(8.5)	(13.4)	(5.7)	(18.2)	(6.9)
Median	9.7	37.9	16.8	17.7	10.3
Chi-Square Value	13.94*	92.74*	0.01	57.01*	6.29*

\* Significant at the  $p < .01$  level

Because of the school differences detected for posttest scores, coded utterances for the two subgroups were analyzed separately. Table 6 shows the mean and median percentages of utterances for each category for the public and private school students separately. Chi-square tests for independence between the public and private school groups were conducted on each of the categories to determine whether the proportion of utterances in each category was significantly different between school types. These

analyses demonstrated that the public school students had a significantly higher proportion of talk related to forethought, motivation, as well as otherwise uncodable utterances. Private school students had a significantly higher proportion of utterances related to performance than did the public school students. Table 6 shows the mean and median percentages by school type, as well as the chi-square values for the tests of independence.

For the second part of this research question, I conducted Pearson correlations between students' gain scores from pretest to posttest, their posttest scores, and the proportions of regulatory utterances in each of the four regulatory categories. Table 7 shows the results of these correlational analyses. Both gain and posttest were significantly related to the proportion of performance, monitoring, and motivation utterances for all audiotaped students. Forethought was not significantly related to either of these learning outcomes. Both performance and monitoring processes were positively correlated with the learning measures, whereas motivation had a negative relation with those measures.

To elucidate the specific processes that resulted in these significant correlations, I conducted analyses on the individual codes within each regulatory category for the audiotaped students. Because of the strong correlations with learning measures, I decided to compare students who were more successful learners (i.e., those who made larger learning gains) to less successful learners (i.e., those who made smaller learning gains) during this task.

Descriptive statistics revealed that the collaborative pairs had a mean gain of 21.2 percentage points, with a standard deviation of 14.3. I selected students who were above

a half standard deviation from the mean to represent the larger gain group (n=16 students), and the students who were below a half standard deviation from the mean as the smaller gain group (n = 16 students). Those who were in between constituted the medium gain group (n = 22 students). These cut points were chosen because they yielded a subsample that represented approximately 60% (n = 32 students) of the audiotaped collaborating students. Of concern was that a more stringent cut point would yield a subsample that was too small to be truly representative of students at the top and bottom of the gain scale.

Table 7

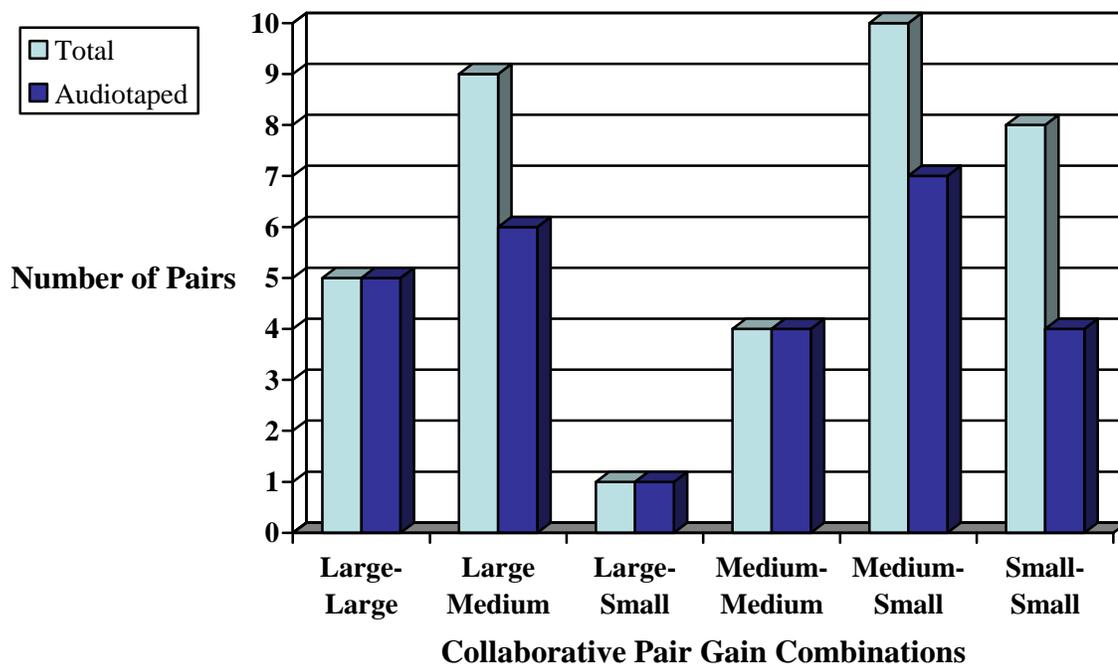
*Correlations between Percentage of Collaborative Regulatory Processes and Gain and Posttest Scores*

	Gain	Posttest	Forethought	Performance	Monitoring	Motivation
Gain	--	.83**	.02	.62**	.35**	-.57**
Posttest		--	.12	.59**	.37**	-.59**
Forethought			--	-.19	.02	-.21
Performance				--	.18	-.80**
Monitoring					--	-.45**
Motivation						--

\*\* Correlation is significant at  $p < .01$

The audiotaped pairs, which were randomly selected during the experimental procedure, were representative of the total collaborative pair sample. Figure 3 shows the distributions of different collaborative pair types, based on the combination of gain score for each student in the pair, as well as number of audiotaped and non-audiotaped for each combination type. For example, the largest number of pairs (n = 10) were those in which one student made a medium gain (within a half a standard deviation from the mean) and

the other made a smaller gain (below half a standard deviation from the mean). Of those 10 pairs, 7 were audiotaped.



*Figure 3.* Number of total and audiotaped collaborative pair gain combination types.

Once the larger and smaller gain students had been identified, raw utterances for each of the regulatory codes were calculated by gain group. Table 8 shows the total frequency (raw counts) and percentages, which reflect the raw frequency divided by overall utterances, by code and group. The median number of overall utterances for students in the larger gain group was 69. For the smaller gain group, the median was 59 utterances.

Table 8

*Raw Frequency and Percentages of Coded Utterances for Larger and Smaller Gain**Groups*

	Larger Gain Group		Smaller Gain Group	
	f	%	f	%
<b>Forethought (Planning)</b>				
Prior Knowledge Activation (PKA)	33	2.9	27	3
Recycle Goal (RG)	5	0.4	11	1.2
Sub-Goal (SG)	30	2.7	20	2.2
Teacher-Set Goal (TSG)	8	0.7	6	0.7
Time and Effort Planning (TEP)	0	0.0	3	0.3
Subtotal	76	6.7	67	7.4
<b>Performance (Strategies)</b>				
Coordination of Information Sources (COIS)	37	3.3	16	1.8
Inference (INF)	35	3.1	29	3.2
Knowledge Elaboration/Explanation (KE)	26	2.3	12	1.3
Memorization (MEM)	15	1.3	1	0.1
Partner Questioning for Procedure (PQP)	25	2.2	37	4.1
Read Notes (RN)	14	1.2	5	0.6
Read Out Loud (RL)	148	13.1	131	14.5
Seeks Consensus (SC)	24	2.1	9	1.0
Search (SEARCH)	14	1.2	6	0.7
Summarize (SUM)	163	14.5	35	3.9
Take Notes (TN)	112	9.9	20	2.2
Subtotal	617	54.8	301	33.3
<b>Monitoring (Reflection)</b>				
Content Evaluation - positive (CE+)	18	1.6	3	0.3
Content Evaluation - negative (CE-)	12	1.1	5	0.6
Expectation of Adequacy of Content (EAC)	18	1.6	13	1.4
Feeling of Knowing - positive (FOK+)	30	2.7	5	0.6
Feeling of Knowing – negative (FOK-)	7	0.6	2	0.2

Table 8, continued

Judgment of Learning – positive (JOL+)	11	1.0	7	0.8
Judgment of Learning – negative (JOL-)	9	0.8	19	2.1
Monitoring Progress Toward Goals (MPTG)	10	0.9	10	1.1
Partner Questioning for Understanding (PQU)	62	5.5	41	4.5
Pretest Monitoring – positive (PM+)	6	0.5	3	0.3
Pretest Monitoring – negative (PM-)	5	0.4	4	0.4
Seeks Affirmation(SA)	13	1.2	7	0.8
Self-Correct (SEC)	1	0.1	3	0.3
Strategy Monitoring (SM)	4	0.4	0	0.0
Time Monitoring (TM)	1	0.1	7	0.8
Subtotal	203	18.0	129	14.3
Motivation				
Off-task (OT)	56	5.0	204	22.6
Positive Feedback (PF)	52	4.6	41	4.5
Negative Feedback (NF)	14	1.2	21	2.3
Positive interest (Int +)	17	1.5	29	3.2
Negative interest (Int -)	4	0.4	8	0.9
Subtotal	143	12.7	303	33.6
Uncodable Utterances (UC)	87	7.7	103	11.4
Total	1126	100.0	903	100.0

Proportions of utterances between the larger and smaller gain groups were compared to determine differences in regulatory processes between the groups. Forethought verbalizations accounted for 6.7% of the total verbalizations of the larger gain group and 7.4% for the smaller gain group. Performance verbalizations accounted for 54.8% and 33.3% for the larger and smaller gain groups, respectively; monitoring verbalizations constituted 18% and 14.3% for the larger and smaller gain groups, respectively; and motivation verbalizations accounted for 12.7% and 33.6% for the larger and smaller gain groups, respectively. Nine of the 37 individually coded processes were

used infrequently by both groups (less than 1% of coded utterances for both groups). They were: stating the *teacher-set goals*, engaging in *time and effort planning*, *searching* the hypermedia environment, stating a *negative feeling of knowing*, *monitoring* related to the *pretest*, *self-correcting*, *monitoring strategy use*, *time monitoring*, and expressing *disinterest*.

Among the codes where one or both groups had a proportion greater than 1%, 11 showed proportions for the two groups that were discrepant by a factor of 2 or greater. In the forethought category, despite a non-significant correlation with learning gains, the smaller gain group engaged in *recycling goals* in working memory at more than twice the proportion of the larger gain group. For the performance processes, the larger gain group engaged in *memorization*, *reading notes*, *seeking consensus*, *summarizing*, and *taking notes* at twice or more the proportion of the smaller gain group. The smaller gain group verbalized nearly twice the proportion of *questioning their partner for procedural* issues than did the larger gain group.

For monitoring processes, the larger gain group expressed a higher proportion (by a factor greater than 3) of positive *content evaluations* and *positive feelings of knowing* than the smaller gain group. The smaller gain group expressed more than twice the proportion of *negative judgments of learning* than those in the larger gain group.

Finally, for motivation-related processes, the smaller gain group had a higher proportion of *off-task* talk (by a factor greater than 4) and, to a lesser degree, expressions of *interest* (by a factor of 2) than the larger gain group.

To highlight differences in regulatory processes between students who made larger gains and those who made smaller gains and how they worked together in groups, I

provide examples and discussion about the discourse between three different collaborating pairs. One purpose of these descriptions is to demonstrate that while comparing proportions of forethought, performance, monitoring, and motivation processes provides information about differences between larger and smaller gain students, looking at the quality of the regulatory processes and the students' interaction helps to more fully describe the different ways these students learned.

The students in each example had similarly low pretest scores, between 6.8% and 14.4 %. However, example 1 is from a pair in which the two students had different gains— one larger-gain student and one smaller-gain student. Example 2 demonstrates two smaller-gain students working together, and example 3 is from two larger-gain students. The excerpts are all taken from discourse occurring within the first 20 minutes of the 30-minute learning session. Codes for the utterances are indicated in brackets. Please refer to Appendix D for the full name and description of the codes.

*Example 1: Large Gain Student with a Smaller Gain Student*

The first example is from a pair in which one student made larger gains and the other partner made smaller gains, as defined by a half a standard deviation above and below the mean, respectively. As Figure 3 demonstrates, this was the only pair in which one student made a larger gain and the other made a smaller gain. In this example, student A is working with student B. Student A is classified as having a smaller gain, with a pretest score of 11.9% and a gain of 11.9% at posttest. Student B is classified as having a larger gain, with a pretest score of 14.4 % and a gain of 33.1%. This pair had a total of 43 utterances between the two, far below the expected median of 128 utterances for a pair where one made a larger gain and the other a smaller gain. Indeed, the students

did not work together to learn the material. Instead, the content of their discourse concerns sharing the hypermedia environment as a shared resource, rather than working together to learn about the circulatory system, as they had been instructed to do.

*Example 1*

Turn	Student	Transcript and Codes
1	<b>Student A:</b>	So, circulatory system [TN] otherwise called the cardiovascular [TN] <i>in humans, the combined function of the heart, blood and blood vessels</i> . [RL] Hey I got that right! [PM+] Damn, I didn't know this part though [PM-] – the immune system and antibodies [TN]
2	<b>Student B:</b>	Oh god.....[TN] ...Ok..I need to write something down before I forget [SM]....look at that, I am genius, how did I know that, I don't know? Common sense maybe?..[FOK+] I am like yeah, man.....[UC]
3	<b>Student A:</b>	Ok, carry away waste. [TN]
4	<b>Student B:</b>	I am good, oh oh oh, now how did I know that [FOK+].....[unintelligible]...ok, what are we....heart and .....[TN] .....[taking notes and reading silently]..
5	<b>Student A:</b>	[taking notes and reading silently].....[TN]
6	<b>Student B:</b>	Where are you now? Can I scroll this up? [SC]
7	<b>Student A:</b>	uh huh [PF]
8	<b>Student B:</b>	And when you are ready, you want to play that? [SC]
9	<b>Student A:</b>	play – oh, see how it works, oh, sure. [PF]..... ..... ..... Three types of cells – oxygen-bearing red blood cells [SUM] – do we have to know all this, goddam! [SA] So, red, white and what now? [TN] what is a blood-clotting plate – whatever?[PQU] ... <i>three types of blood vessels</i> [RL] I can't be bothered with this anymore. [INT-]
10	<b>Student B:</b>	Ready? [starts video] ..... [FW: you have about 15 minutes left]

- 
- 11 **Student A:** What a bore [INT-]
- 12 **Student B:** Do you want to go to operation and function? [SC]
- 13 **Student A:** I am tired, I don't feel like doing this anymore – got boring..[INT-] Oh my gosh.....  
.....
- 14 **Student B:** How much time do we have? [TM] .....
- 15 **Student A:** If you need to like scroll down, go ahead.....  
You can keep going.....[SC]
- 

Note: slashes (/) indicate beginning and end of an utterance; italicized words indicate reading aloud; non-italicized words are other utterances;

[TN] = Taking notes

[RL] = Reading out loud

[PM+] = Pretest monitoring, positive

[PM-] = Pretest monitoring, negative

[SM] = Strategy monitoring

[UC] = Uncodable

[FOK+] = Feeling of knowing, positive

[SC] = Seeks consensus

[PF] = Positive Feedback

[SUM] = Summarizing

[PQU] = Questioning Partner for Understanding

[INT-] = Interest, negative

[TM] = Time monitoring

As can be seen from the example, there was very little reading out loud of any of the text. Instead, the students read silently to themselves and took their own notes with little collaboration. Turns 6 through 9, 10, 12, and 15 were discourse about sharing the computer and where to navigate next. In turn 9, student A, the smaller gain student, started to become disinterested in the task, and every turn of his thereafter expressed this disinterest. Student B worked by himself, talking only to his partner about where to move in the hypermedia environment (turns 6, 8, 10, and 12). In sum, these two students did not work collaboratively to learn, but rather treated the task as a sharing-of-resource task.

Further, the student who made smaller gains showed evidence of being uninterested in the task several times in this excerpt.

*Example 2: Two Students with Smaller Gains*

In the second example, both students made smaller gains. Student C had a pretest score of 6.8% and actually scored slightly below this on the posttest, for no overall gain. Similarly, student D scored 11% at pretest and made no overall gains on her posttest. This pair had a total of 94 utterances between the two, which was fewer than would be expected based on a median utterance for the smaller gain students of 59 (or 108 for a pair).

*Example 2*

Turn	Student	Transcript and Codes
1	<b>Student C:</b>	<i>heart, blood antibodies and immune system and plasma</i> [RL]...I hope this test is multiple choice [posttest].[UC]
2	<b>Student D:</b>	Do we have to read that?..[PQP]
3	<b>Student C:</b>	<i>Three types of blood vessels form a complex network of tubes throughout the body. Arteries carry blood away from the heart, and veins carry it toward the heart.</i> [RL]
4	<b>Student D:</b>	<i>Capillaries are the tiny links between the arteries and the veins where oxygen and nutrients diffuse to body tissues. The inner layer of blood vessels is lined with endothelial cells that create a smooth passage for the transit of blood. This inner layer is surrounded by connective tissue and smooth muscle that enable the blood vessel to expand or contract. Blood vessels expand during exercise to meet the increased demand for blood and to cool the body. Blood vessels contract after an injury to reduce bleeding and also to conserve body heat.</i> [RL]
5	<b>Student C:</b>	So they contract...hm....[SUM]
6	<b>Student D:</b>	<i>Arteries have thicker walls than veins to withstand the pressure of blood being pumped from the heart. Blood in the veins is at a lower pressure, so veins have one-way valves to prevent</i>

---

*blood from flowing backwards away from the heart. Capillaries, the smallest of blood vessels, are only visible by microscope—ten capillaries lying side by side are barely as thick as a human hair. [RL] Geez! [INT+]*

- 7 **Student C:** Really?! [INT+] *If all the arteries, veins, and capillaries in the human body were placed end to end, the total length would equal more than 100,000 km (more than 60,000 mi)—they could stretch around the earth nearly two and a half times[RL].... Kind of amazing! That is kind of weird.[INT+]*
- 8 **Student D:** *The arteries, veins, and capillaries are divided into two systems of circulation: systemic and pulmonary. The systemic circulation carries oxygenated blood from the heart to all the tissues in the body except the lungs and returns deoxygenated blood carrying waste products, such as carbon dioxide, back to the heart. The pulmonary circulation carries this spent blood from the heart to the lungs. In the lungs, the blood releases its carbon dioxide and absorbs oxygen. The oxygenated blood then returns to the heart before transferring to the systemic circulation. [RL]*
- 9 **Student C:** That is a lot of information in a small time period – that is why I think it is going to be hard to remember. [JOL-]
- 10 **Student D:** I think we should stop [TEP]
- 11 **Student C:** Let's go back and – [SC] oh look, that changes up here, too – pink[UC]
- 12 **Student D:** What is she doing with the paper? [referring to experimenter].....I am really hungry![OT]
- 13 **Student C:** Do you eat breakfast?[OT]
- 14 **Student D:** No – do you?[OT]
- 15 **Student C:** Every day - I would not be able to make it through the day if I didn't.[OT]
- 16 **Student D:** When I eat breakfast, I am usually more hungry.[OT]
- 17 **Student C:** That is weird. Like, sometimes I don't feel hungry when I wake up, but I still eat it – like I eat a pop tart. I eat – I snack in my car, and then when I get home, dinner is ready, then I eat later – I eat all the time.[OT]
-

- 
- 18 **Student D:** Like, last week I spent \$300 on my dad's credit card on food![OT]
- 19 **Student C:** Sheez!..That's bad.....[OT]
- 20 **Student D:** What are we supposed to do?[PQP] I don't want to read all that...[INT-]
- 

Note: slashes (/) indicate beginning and end of an utterance; italicized words indicate reading aloud; non-italicized words are other utterances.

[RL] = Reading out loud

[UC] = Uncodable

[PQP] = Questioning partner for procedure

[SUM] = Summarizing

[INT+] = Interest, positive

[JOL-] = Judgment of learning, negative

[TEP] = Time and effort planning

[SC] = Seeks consensus

[OT] = Off task

[INT-] = Interest, negative

Relative to the students in the first example, students C and D worked together more. They each took turns reading out loud (turns 3, 4, 6, 7, and 8). However, there was little discussion or summarization that occurred between these episodes. The students did not elaborate on what they are reading or make inferences, and they did not seem to monitor their understanding or their progress in the task. And, although each expressed interest during the first part of the excerpt (turns 6 and 7), their interest was directed toward specific small pieces of information rather than the task as a whole. Indeed, in turn 9, student C expressed that he was finding the task difficult. In turns 12 through 19, the students discussed off-task topics. In fact, of their 94 utterances, 16 utterances (approximately 17%) were off-task talk, which is consistent with the relatively high proportion of off-task talk for all smaller gain students (23%; Table 7).

In sum, the students in this pair made no gains from pretest to posttest. They did work together in a collaborative manner, but they engaged in surface-level processes only, such as reading out loud, with little effort put into other comprehension or

monitoring strategies. As well, 17% of their utterances were off-task talk, unrelated to learning about the circulatory system.

*Example 3: Two Larger-Gain Students Working Together*

The third example exemplifies two larger-gain students working together. Student E had a pretest score of 13.6% and gained 49.2% on her posttest. Student F scored 11.9% on her pretest and made a gain of 30.5 % on her posttest. This pair had a total of 117 utterances between the two, which was fewer than would be expected based on a median utterance for the larger gain students of 69 (or 138 for a pair).

*Example 3*

Turn	Student	Transcript and Codes
1	<b>Student E:</b>	Alright.. ew – ok, so <i>the blood vessels, heart valves, heart – the human heart is a hollow-pear shaped organ about the size of a fist [continues reading aloud]</i> [RL] Ok, so it enters through there, and it also enters through there [COIS]
2	<b>Student F:</b>	Right [PF]
3	<b>Student E:</b>	and it collects there and once it is there, [SUM] <i>the atrium fills it contracts and the blood passes through the tri</i> [RL]– through here. [COIS]
4	<b>Student F:</b>	Uh huh...[PF]
5	<b>Student E:</b>	tricuspid valve to the right [SUM]
6	<b>Student F:</b>	There [COIS]
7	<b>Student E:</b>	where is the right atrium – oh [PQU]
8	<b>Student F:</b>	Here [COIS]
9	<b>Student E:</b>	<i>right ventricle becomes full and starts to contract and the mitral valve closes to prevent the blood from [continues reading aloud]</i> [RL]
10	<b>Student F:</b>	<i>[reading aloud with partner]</i> [RL] pulmonary artery – right there [COIS]

- 
- 11 **Student E:** *Which carries blood to the lungs to pick up fresh oxygen. When blood exits the right ventricle [RL] – which is right there – [COIS]*
- 12 **Student F:** right there [COIS]
- 13 **Student E:** *the pulmonary valve shuts preventing blood from moving back in the right ventricle. Blood returning from the lungs to the heart collects in the left atrium... [RL] so I am guessing that is blood with oxygen in it... [INF]*
- 14 **Student F:** then it returns here [COIS]
- 15 **Student E:** So then it returns into there [COIS]
- 16 **Student F:** Right [PF]
- 17 **Student E:** so I guess that blood has oxygen in it – [INF – counted above] *so this chamber contracts, blood flows through a valve into the left ventricle [RL]– right there [COIS]*
- 18 **Student F:** Right [PF]
- 19 **Student E:** *and then the mitral valve between the two closes. In the final phase of blood flow through the heart, the left ventricle contracts and forces blood into the aorta. [RL] So its that valve [COIS]*
- 20 **Student F:** yep..[PF]
- 21 **Student E:** *Um, after the blood in the left ventricle has been forced out, the mitral valve closes and aortic at the opening of the aorta. [RL] So it goes in there, then in there, hangs out in there, then it like starts pumping and it goes in there and then it pumps it up there here through the pulmonary artery, goes into the lungs, gets some oxygen, comes back in here and then goes out to the rest of the body. [SUM]*
- 22 **Student F:** uh huh [PF]
- 23 **Student E:** Now let's look at the heart valve...[SG] um.....*The heart's pulmonary artery and aortic valve [reading aloud] [RL] so that is like what it was talking about with the heart valve [FOK+]*
-

- 
- 24 **Student F:** right, [PF] closes and comes back in.[SUM]
- 25 **Student E:** *[reading aloud].....as the heart relaxes between one heartbeat and the next, blood pressure falls [reading aloud]....[RL]*
- 26 **Student F:** closes to stop blood. [SUM]
- 27 **Student E:** *blood vessels.....blood vessels circulate blood through the body – three major types of blood vessels are arteries veins and capillaries. [RL] Arteries are the really big ones, veins are the medium and capillaries are the smaller ones. [SUM] Arteries carry blood away from the heart while veins carry blood toward the heart. Capillaries - [RL] so, arteries carry blood away from the heart. [SUM] So arteries are the ones that carry it to the brain and the rest of your body [KE]*
- 28 **Student F:** Right [PF]
- 

Note: slashes (/) indicate beginning and end of an utterance; italicized words indicate reading aloud; non-italicized words are other utterances.

[RL] = Reading out loud

[COIS] = Coordinate informational sources

[PF] = Positive feedback

[SUM] = Summarizing

[PQU] = Questioning Partner for Understanding

[INF] = Inference

[SG] = Sub-goal

[FOK+] = Feeling of knowing, positive

[KE] = Knowledge elaboration

Student E and student F collaborated well together. They not only read out loud to each other, but they constructed knowledge together, building upon each others' verbalizations. This is most clear in turns 5 through 15, during which the pair read text that describes blood flow through the heart while looking at an accompanying diagram. Student F also gave positive feedback to her partner on many occasions (turns 2, 4, 16, 18, 20, 22, 24 and 28), which served to encourage her partner to continue reading and thinking out loud about what she was learning. In turn 13, student E made an inference that is critical to fully understanding the double-loop nature of blood-flow through the

heart and lungs. In turn 27, she made a knowledge elaboration demonstrating that she connected what she was learning with prior knowledge. Finally, both students engaged in active learning strategies such as summarization and coordination of information sources.

In sum, this pair of students, both of whom made larger gains from pretest to posttest, worked together collaboratively. They used active learning strategies, constructed knowledge together, and gave positive feedback. Further, they identified information that was critical to understanding circulation conceptually.

### *Research Question 3*

Does answering a reasoning question while learning about a complex science topic facilitate high-school students' conceptual-knowledge learning? Visual inspection of the data revealed a potential effect of school type on pretest and posttest scores. T-tests between the private and public schools on pretest and on posttest scores revealed no significant differences on the pretest scores between the two school type groups,  $t(131) = -0.98, p > .05$ , but significant differences on the posttest scores,  $t(131) = 6.19, p < .05$ , between the two groups. Despite school type differences, the variability within each school type on posttest was nearly equivalent to the variability between them. As such, school type was not added as an effect in the subsequent analysis of covariance.

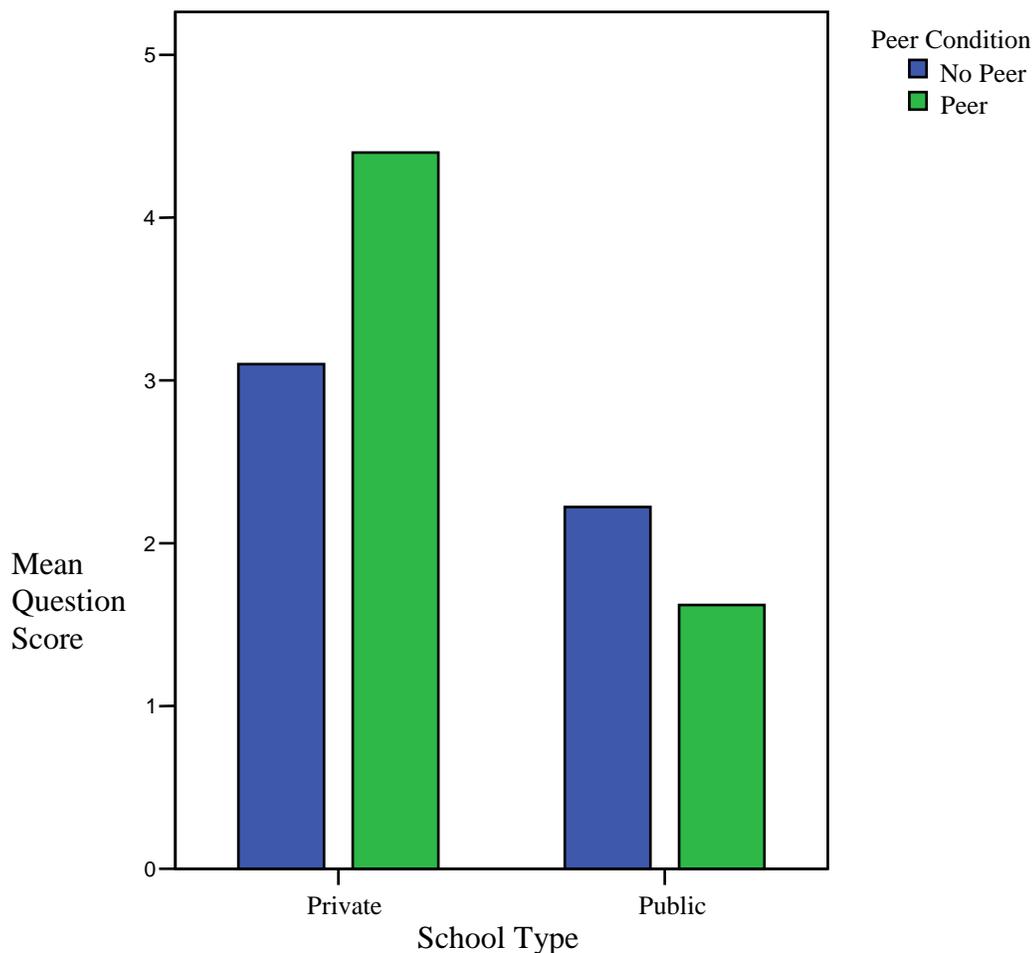
A 2 X 2 analysis of covariance, with pretest as the covariate and posttest as the dependent variable, revealed a non-significant main effect for questioning condition,  $F(1, 128) = 0.14, p > .05$ , indicating those learning with a reasoning question did not have significantly higher adjusted posttest scores than those learning without a reasoning question. See Table 5 for a summary of the analysis.

#### *Research Question 4*

Does working with a peer facilitate explanatory reasoning associated with answering a reasoning question to a greater degree than working alone? Inspection of the data on question response scores also indicated a difference between school types. A significant t-test between School Type on question score,  $t(62) = 4.26, p < .05$ , revealed that students at the private schools scored higher on average ( $M = 3.88, SD = 1.64$ ) than those at the public school ( $M = 1.90, SD = 1.91$ ) on question response scores. A 2 (Peer Learning Condition) X 2 (School Type) analysis of variance on question response scores from those who were in the Question Condition demonstrated no main effect for Peer Learning Condition,  $F(1, 61) = 0.54, p > .05$ . However, there was a significant interaction effect between School Type and Peer Learning Condition,  $F(1, 61) = 13.28, p < .05$ . Figure 4 demonstrates this interaction graphically.

Simple effect analyses for the private school type revealed a significant difference in question response score,  $t(24) = 2.09, p < .05$ , between those who worked with a partner and those who worked individually. This analysis demonstrated that students in the private school who worked with a peer had higher average question response scores ( $M = 4.40, SD = 1.40$ ) than those who worked individually ( $M = 3.10, SD = 1.73$ ).

Simple effect analyses for the public school type revealed a non-significant difference in question response score,  $t(37) = -0.89, p > .05$ , between those who worked with a partner and those who worked individually. This analysis demonstrated that students at the public school who worked with a peer ( $M = 1.62, SD = 1.71$ ) did not have significantly higher average question response scores than those who worked individually ( $M = 2.22, SD = 2.13$ ).



*Figure 4.* Interaction between school type and peer learning condition on mean question response score.

#### *Research Question 5*

Do collaborative dyads answering a reasoning question while learning engage in greater instances of utterances related to forethought, performance, reflection, or motivation than collaborative dyads not answering a reasoning question? An independent groups Mann-Whitney  $U$  test of ranks for non-parametric data revealed no significant differences in mean rank of students in the peer with question condition compared to

those in the peer with no question condition, based on number of utterances related to forethought ( $U = 272, Z = -1.22, p > .05$ ), performance ( $U = 309, Z = -0.56, p > .05$ ), monitoring ( $U = 322, Z = -0.32, p > .05$ ), and motivation ( $U = 237.5, Z = -1.84, p > .05$ ).

This analysis was also conducted on public and private school student groups separately, because significant differences in question response score were found between private school students working with peers and those working individually (see research question 4). A Mann-Whitney  $U$  test on the private school students working with a peer found significant differences in mean rank of students based on utterances of forethought ( $U = 36.5, Z = -2.60, p < .05$ ), indicating students in the question condition engaged in a greater frequency of utterances related to forethought. No significant differences between the groups was found for utterances related to performance ( $U = 58, Z = -1.54, p > .05$ ), monitoring ( $U = 58, Z = -1.54, p > .05$ ), or motivation ( $U = 88, Z = -0.10, p > .05$ ). A Mann-Whitney  $U$  test on the public school students revealed no significant difference in mean rank of students in number utterances related to performance ( $U = 32, Z = -2.53, p < .05$ ), forethought ( $U = 52.5, Z = -1.45, p < .05$ ), and monitoring ( $U = 49, Z = -1.63, p > .05$ ). However, students in the question condition ranked significantly higher in their utterances related to motivation ( $U = 25.5, Z = -2.87, P < .05$ ) compared to their classmates who did not have a reasoning question. Table 9 shows the mean rank values and results of the Mann-Whitney  $U$  test for each group of students.

Table 9

*Sample Size, Mean Ranks, and Mann-Whitney U Value for Proportions of Utterances Related to Forethought, Performance, Monitoring, and Motivation*

	Question		No Question		U Value	P
	N	Mean Rank	N	Mean Rank		
Overall	20		34			
Forethought		30.90		25.50	272.00	.22
Performance		29.05		26.59	309.00	.58
Monitoring		28.40		26.97	322.00	.75
Motivation		32.63		24.49	237.50	.07
Private	10		18			
Forethought		19.85		11.53	36.50	.01*
Performance		17.70		12.72	58.00	.13
Monitoring		17.70		12.72	58.00	.12
Motivation		14.30		14.61	88.00	.92
Public	10		16			
Forethought		11.55		14.72	60.50	.30
Performance		11.90		14.50	64.00	.42
Monitoring		11.20		14.94	57.00	.24
Motivation		19.25		9.91	22.50	.00*

\* Significant at  $p < .05$

## CHAPTER 5: DISCUSSION

The purpose of this dissertation was to explore the role of collaboration and of reasoning questions in helping students develop conceptual knowledge about a complex system, the circulatory system, while they learned using a hypermedia encyclopedia. To increase ecological validity, the study was conducted in intact classrooms with student participants from both public and private high schools. In this chapter I discuss the three major questions guiding this research, summarize and synthesize several major themes that emerged, and describe potential implications for teaching and learning, issues for future research, and limitations of the study.

Before addressing the results relative to the guiding research questions, it is necessary to discuss the unexpected effect of school type that arose in analysis of the data. In particular, I found that students from the private schools had, on average and across conditions, significantly higher average posttest scores than those who attended public school. I intentionally drew participants from both public and private schools so that the sample of students in the study would be more representative of the population of high-school students than a sample drawn exclusively from a convenience sample of private schools. The finding was unexpected because students from both school types had similarly low prior knowledge about the topic, which was reflected in their pretest scores. The public high school is considered one of the best in its state and consistently scores above average on state-mandated standardized tests, with most of its graduates going on to college. There was no overt indication that the task itself (i.e., working with a peer to research a particular topic using hypermedia) was novel or foreign for any of the students at any of the schools. The teachers at both schools indicated that the activity would be an

appropriate one for their students; that it fit within their curriculum; and that students often worked together on short projects such as this.

One obvious difference between the schools was the size of the instructional classes. The size of the classes in the private school averaged 12 students with one teacher, while the average size of the classes in the public school was 25 students with one teacher. Indeed, my observation was that the classroom environments were quite different between the two school types. At both private schools, the classrooms appeared less chaotic and less noisy than the classrooms at the public school. The public-school students appeared to socialize with their fellow students (not about the task) at any opportunity much more than those at the private school, who appeared to be more attentive to me and to the task from the start of the class. As such, I had to expend more effort at the beginning of the public-school classes getting students organized and focused on the task as compared to the students at the private school. Research on class size has identified differences in student behavior in small classes as compared to larger classes. For example, it has been demonstrated that students in small classes tend to be more active participants in learning, engage in less “social loafing” and have a greater sense of involvement and responsibility in the classroom than those in larger classes (Finn, Pannozzo, & Achilles, 2003). Further, small-class environments tend to encourage more peer-peer and student-teacher interaction (Finn et al., 2003; Rice, 1999). It is possible that the private-school students, used to the atmosphere of a small-class learning environment, approached the task differently based on their experiences in such an environment. This experience may have included more occasions for open-ended learning with ill-structured tasks, a greater emphasis on self-directed learning, and a greater emphasis on active

involvement in classroom activities. As such, the private-school students may have focused and stayed on task and been more invested overall in the task than those students at the public school, largely because of their school experience.

Even though the differential performance between school types was unexpected, it may help illuminate several other results in this study by highlighting the importance of academic experience or classroom climate in successful collaborative learning. I will revisit this idea in the conclusion of this section, with a discussion of how this experience may affect factors related to success in peer collaboration.

### Peer Collaboration

The first major question guiding this study was whether working with a peer would facilitate students' conceptual-knowledge learning to a greater degree than working alone. For this study, working with a peer did not, on average, provide any advantage to learning conceptual knowledge compared to learning individually. This result can be attributed to the fact that there was wide variability across pairs in learning gains. In contrast, within pairs, students usually had gain scores (from pretest to posttest) that were similar to that of their partner. In fact, there was only one group out of the 37 with a large discrepancy; that is, in which one student demonstrated large gains and her partner had small gains. The other groups were nearly evenly divided between having both students in the same gain group and having students in adjacent (e.g., medium and small) groups (see Figure 3). Similarly, Jeong and Chi (2007) found that collaborating pairs constructed shared common knowledge – inaccurate as well as accurate – during collaborative text comprehension. This provides some evidence that peers working together influenced each others' learning, for better or for worse, and supports the contention that collaborating students build a joint-problem space and may create a

“shared conception of the problem” (Roschelle & Teasley, 1995, p. 70). In this study, the ways in which student pairs collaborated together were variable in both approach, conception of the problem, and in learning success.

To better determine how students’ learning behaviors may have contributed to their variable learning success, I looked at the relation between students’ regulatory verbalizations and their learning gains. The proportions of performance, monitoring, and motivation regulatory processes were significantly correlated with posttest scores and learning gains, indicating a relation between these processes and learning. To elucidate which specific processes were contributing to those significant correlations, I compared the proportion of verbalizations of the specific regulatory processes of students who made larger learning gains to those who made smaller learning gains. There were a number of regulatory processes that the larger-gain students engaged in to a greater degree (by a factor of two or more) than the smaller-gain students. These included *memorizing*, *reading notes*, *seeking consensus*, *summarizing*, *taking notes*, *evaluating content*, and *expressing feelings of knowing*. The smaller gain group engaged in *recycling goals* in working memory, *negative judgments of learning*, *off task behavior*, and *positive interest* to a greater degree than those in the larger gain group.

To determine how these specific processes played out in collaborative groups, I presented examples in Chapter 4 that illustrated three different approaches to the collaborative task. These examples echo the findings of Forman and Cazden (1994), who identified three profiles of collaboration between student pairs working on a mathematics problem-solving task: parallel, associative, and cooperative. Students who work in parallel share resources, but they do not monitor each others’ actions or inform each other

of thoughts and actions, similar to example 1 (see p. 115). In this pair, the students did not work collaboratively to learn, but rather appeared to perceive the task as individuals sharing a computer resource. Students who work associatively share information, but make no effort to coordinate roles, similar to example 2 (see p. 117). In the second example, the students interacted more than those in example one, but the interaction consisted of surface-level behaviors such as reading out loud, and there was little discussion between the two related to understanding or synthesizing the content. Lastly, cooperative dyads are pairs who monitor each others' progress and roles and share information, similar to example 3 (see p. 120). In the current study, the third pair worked together to build a shared understanding of the circulatory system.

These examples demonstrate that the way students interacted may have been related to the learning success of the students involved. The students in example 1 did not collaborate, and this may have contributed to their quite different posttest scores, despite having similar pretest scores. Further, even though both students engaged in a high degree of taking notes, a behavior associated with the larger gain group, these notes were done independent of each other with no discussion ensuing between reading and taking notes. In this example, each student regulated his own learning independent of his partner. One student (student A) was likely more successful at doing so than his partner (student B), which bore out in their disparate posttest scores.

In example 2, neither student made gains from pretest to posttest. While it appeared they worked together on the task, they did not employ strategy and monitoring processes associated with the larger gain group, such as summarizing, taking notes, evaluating content, or expressing feelings of knowing. They approached the task at a

surface-level—reading out loud to each other with little discussion of what they were learning. They expressed interest, as was characteristic of smaller gain students overall, but the topics about which they expressed interest were largely tangential to the overall conceptual understanding of the circulatory system that was the goal of this task. One possibility is that their attention to these tangential (or seductive) details may have interfered with their learning (Garner, Gillingham, & White, 1989; Harp & Mayer, 1998). This example might help explain why smaller gain students had a higher proportion of positive interest statements compared to larger gain students, despite this seemingly counterintuitive finding. Finally, the students in this pair engaged in a considerable amount of off-task talk, as was characteristic of all the smaller-gain students, and which was not conducive to learning the topic at hand.

Example 3 described two larger-gain students. They worked together collaboratively, engaging in strategies that were characteristic of larger gain students, such as summarizing. They also engaged in strategies associated with active learning, such as knowledge elaboration and inference generation. Interestingly, proportions of processes such as knowledge elaboration and making inferences did not differ dramatically between the larger and smaller gain groups overall. Not captured in the proportional analyses, however, was the quality of the strategy used. In example 3, the students made an inference about information critical to understanding the flow of the human circulatory system and could therefore be considered a high-quality inference within the context of this task. It is possible that smaller-gain students made inferences, but not ones that were instrumental in developing their conceptual understanding.

These results are similar to those of prior research on peer collaboration in science. Several of the reviewed studies found evidence that successful peer collaborators: actively sought information and were more metacognitive than those who were not as successful (Kneser & Ploetzner, 2001); engaged in more problem-centered activities than low-gain peer learners who engaged in surface-level activities (Chan, 2001); and used more “interpretive talk” involving planning, predicting, and using strategies than those who used “descriptive task” focused on procedure and describing evidence and who were less successful learners (Teasley, 1995). The results of this study serve to bolster the evidence that successful peer collaborators engage in deeper-level processes that move beyond procedural and other surface-level activities.

Prior research in peer collaboration has often focused on students ability levels, with some success providing evidence that more “able” peers can help less “able” peers without detriment to their own learning (e.g., Lou et al., 2000). This study did not measure students’ abilities, and students who worked with a peer were randomly assigned within gender to a partner. Are the results from this study related to the ability levels of the students collaborating? In other words, were successful pairs those that had the appropriate mix of abilities? Carter and Jones (1994) found that when two low-ability (as determined by low CAT scores) peers worked together, they were more off-task than when a low-ability and high-ability peer worked together. Further, Saleh et al. (2005) found that students engaged in more elaboration when they were in mixed (as opposed to heterogeneous) pairs. However, the ability literature defines “ability” in many ways that could be redefined as achievement, motivation, or prior knowledge. As such, the finding

that ability level affect peer collaboration may actually be masking other individual differences that need to be elucidated.

On the surface, varied levels of engagement with the task appear to explain much of the variability in the peer interaction in this study. For example, a number of pairs had low engagement in the task, as evidenced by a large proportion of off-task talk, such as in example 2, while others stayed on task for the duration of the time, as with the pair in example 3. However, there were a number of pairs that appeared motivated to learn and to complete the task. Nonetheless, they were not very successful at learning the critical content necessary to demonstrate large learning gains. These student pairs appeared to perceive the fundamental aim of the task differently than those who were more successful at conceptual-knowledge learning. A number of less-successful pairs tended to focus on learning declarative knowledge only, to the exclusion of putting this information together to construct an understanding of the cycle of blood flow. An explanation for this could be the low prior knowledge of the less successful students. However, in this study, pretest scores were similarly low for all but a few students. In an open-ended learning situation, a student's learning outcome will be strongly related to what the student perceives the task to be and what he or she aims at learning (e.g., Grossen, 1994; Marton & Säljö, 1976) and similarly so with collaborating pairs. As Pintrich (2000) explained, task perception is an important contextual component of self-regulated learning; it is not something the student can self-regulate, but it does determine the goals, strategies, and level of engagement that students regulate as they learn. By extension, the mutual task perception built by students collaborating may guide the ways in which they collaboratively regulate their learning. Students' task perception may have influenced

how students navigated through the hypermedia environment and on the ways in which they interacted with their peers and the content. The data from this study suggest that students approached the task differently based on their perception of what the task entailed.

Another plausible explanation for why no significant differences in learning appeared between pairs and individuals is that the task was timed—30 minutes for both individuals and pairs. It is possible that the pairs spent less actual time engaging with the content than did individual learners because some time needed to be spent interacting with partners, negotiating navigating through the environment, and making decisions about next steps in the task. In which case, any advantage conferred by working with a partner may have been offset by the time constraint, a phenomenon in group work that has been called *process loss* (Steiner, 1972). Prior research supports this contention. For example, in their research on peer collaboration, Bianchini (1997) and Kumplainen et al. (2001) found that peer collaborators spent much of their time on procedural and organizational issues. Given these considerations, it is worthwhile to note that peer learners and individual learners had statistically indistinguishable average adjusted posttest scores. If the time limit had been longer or indefinite, the result might have been different. However, it did appear that many pairs felt they were finished with the task before the time limit ended, indicating potentially ineffective collaborative metacognitive monitoring relative to their learning and the goals of the task. This observation serves to further highlight the possible role that task perception played in how students undertook the task.

As has been demonstrated in the literature on peer collaboration, there are often social factors that can influence the outcome of peer collaboration. For example, Cohen (1984) found that the social status of the peers working together influenced their interactions and outcomes of the interaction. Low status students talked less and asked for more help, which was often ignored by the higher status partners (Cohen, 1984). Measurement of these factors was beyond the scope of this study, but their potential influence should not be discounted. It is possible that these differences help explain some of the variability in outcomes.

In sum, a number of viable explanations exist for why peer learners did not demonstrate significantly larger learning gains than students working individually. There was high variability in the peer learners, in the extent to which they engaged with the task and how they perceived the aim of the task. These, along with a time limit on the task and possible social influences, may have been related to high variability in outcomes.

### Reasoning Questions

The second major question guiding this study was whether students who were provided with a reasoning question would demonstrate greater conceptual-knowledge learning than those without a reasoning question. The results of this study indicated that the reasoning question did not help students learn more conceptual knowledge than those without the question.

Several potential explanations exist for why the reasoning question did not help students' conceptual-knowledge learning in this circumstance, despite evidence in prior questioning research showing a connection between higher-order questions and learning (Hamaker, 1986). First, the relatively short time given to students to learn using a hypermedia environment may have thwarted those who were given questions to answer,

thus masking any effect that a longer session might have yielded. In particular, transcripts of pairs answering the reasoning question revealed that many spent time searching for the answer to the question, expecting a verbatim response from the environment. Only after spending some time searching did many pairs start making inferences and attempting to answer the question. Without think-aloud data from the individual learners, one can only conjecture that similar behavior was present with them.

This behavior may have presented a problem because what should have been a constructive, inferential learning exercise was treated by many as a simple search task, involving lower cognitive processing of information. Students may have perceived the question answering task as a finding-of-information task, rather than an understanding, learning, and remembering task. The relatively low average question response scores provide evidence that many students did not find or learn the appropriate information to answer the question fully, and thus have a better understanding of blood flow through the heart and body.

Coupled with an incorrect perception of the task, the selective-attention effect may have been occurring in this circumstance, with students attending to information that was not important for demonstrating conceptual understanding, despite the intention of the reasoning question. The selective-attention hypothesis, which has support in the literature (e.g., André, 1979; Holliday & McGuire, 1992; Reynolds & Anderson, 1982) posits that one role of questions is to direct learners' attention to important information that is covered on a subsequent measure of learning. In this study, the reasoning question was considered related to the flow item on the posttest, because answering the question correctly entailed understanding the blood flow through separate sides of the heart and

lungs. Students did not receive feedback about the correctness or completeness of their answers, so it was up to them to gauge when they thought they had completed the task of answering the question.

Another explanation can be found in prior research. In particular, Wang and Andre (1991) found that a pretest directed students' attention to the material they needed to learn, masking any effect of a question to direct their attention; a similar situation may have occurred with this study, whereby students who had a reasoning question were already selectively attending to the material presented on the pretest. However, students' ability to attend to the appropriate information appeared inadequate, or they would have had higher posttest scores. One reason students may not have attended to the appropriate information is that they had relatively low prior knowledge about the topic. Research on student questioning has revealed a positive relation between prior knowledge and sophistication of question asked by students (e.g., Miyake & Norman, 1979; Scardamalia & Bereiter, 1992; Taboada & Guthrie, 2006). By a similar mechanism, students with low prior knowledge may not be able to adequately answer a higher-order question asked of them, and more so if they do not perceive the aim of the question accurately. The students in this study may not have had enough knowledge to know what they did not know, where they could find what they wanted to know, or to gauge what might be most important to learn (Alexander & Jetton, 1996). Students' lack of prior knowledge also likely meant their search process was inefficient; without a developed schema of the topic, students using hypermedia often do not utilize links and nodes well (Dillon & Jobst, 2005). The time students spent searching the environment was underutilized

learning time that could have been otherwise spent reading and understanding rather than searching.

It is also possible that there are developmental differences in how questioning affects student learning. The majority of research on adjunct questioning and elaborative interrogation has used college students as participants and generalized across age levels (e.g., Hamaker, 1986; Ozunger & Guthrie, 2004; Pressley et al., 1988). However, Van der Broek, Tzeng, Risdén, Trabasso, and Basche (2001) found that fourth-, seventh-, and tenth-grade students who were given embedded inferential adjunct questions had less recall than students in the control groups in each respective grade who did not have embedded questions. However, college students who were provided with the questions had greater recall than a control group of college students. Van der Broek et al. (2001) attributed this difference to differences in reading skills at these developmental levels. It is possible that the results from this study might have been different with a more skilled sample of readers, such as those at the college level.

#### Collaborating on Question Answering

The final major question guiding this study was whether those working with a peer would demonstrate better explanatory reasoning in answering the reasoning question compared to those working individually. The results for this inquiry revealed that overall, individual and peer learners were indistinguishable in question response score. However, students from the private schools who worked with a peer to answer the reasoning question had higher quality answers to the reasoning question than those private-school students who worked individually to answer the reasoning question and compared to individual and peer public-school learners.

In comparing the public and private-school peer learners who had a reasoning question, it appeared that the private-school students had a perception of the task that was more in line with what I had intended. The discrepancy in answer scores between public- and private-school students also suggests that public-school students are not as used to answering reasoning questions such as the one provided in this task—either alone or in small groups. The average score of question response for public school students working with a peer was 1.62 out of 5, while for private school students it was 4.40 out of 5. The full answer to the reasoning question was not readily available within the environment. Rather, students needed to pull together several pieces of information and make inferences to fully answer the question. Many of the students (public and private) gave some answers that made sense with prior knowledge (i.e., gravity). As such, it is possible that students' expectations (based on their school experience), and definitions of a task dictate the ways in which they go about a task. In this case, a surface approach led to students stopping once one or two obvious answers were found. And, for this sample of students, this occurred more frequently with public-school students.

In comparing the frequency of collaborative self-regulatory talk between pairs who did and did not have a reasoning question, results showed that private school students in the questioning condition had a significantly higher average frequency of talk categorized under forethought compared to their classmates without a question. A closer look at the particular utterances within each of these groups reveals some explanations for these results. For the private-school students, the relatively high frequency of forethought talk was due to more prior knowledge activation and more goal-related (main goals, sub-goals) talk in the question condition as compared to those in the no-question

condition. It is apparent from these results that, for the private school students working with a peer, having the question to answer stimulated students to activate their prior knowledge verbally to their partner, as well as encouraged goal-setting around answering the question. This result provides support for the idea that questioning a learner may facilitate active construction and integration of knowledge into prior knowledge (Brown et al., 1983; Craig et al., 2006). Further, providing the question stimulated these students to engage in collaborative regulatory behavior in the form of setting goals.

Why, then, did the public school students not show such a result? One explanation arises from the verbal discourse data. Public-school peer learners in the questioning condition had significantly more off-task talk (categorized as motivation-related behaviors) compared to their classmates without a question. The average proportion of off-task talk for public-school students who worked with a peer and who had a reasoning question was 26.2%, while private-school students in the same condition had a proportion of 7.0% off task -talk. By comparison, public-school and private-school students working with a peer who did not have a reasoning question had an average proportion of 6.6% and 3.4% off-task talk, respectively. This high proportion of off-task talk, which came predominantly from three of the five audiotaped public-school pairs that had a reasoning question, may indicate a lack of engagement in the task.

Closer inspection of the three pairs who had high off-task talk reveal that they all attempted to answer the question, but for two of the three pairs the students went off task once they felt satisfied with their answer to the question, even though there were more sophisticated answers they could have provided. For these pairs, the obvious answers were enough, and they did not attempt to learn more and develop more complex answers.

These students appeared to define this task narrowly. The remaining pair engaged in off-task talk off and on throughout the task and appeared only mildly engaged in the task from the very beginning (e.g., “what are we doing?”). This type of behavior may indicate disengagement with a too-difficult or confusing task. In contrast, the private-school students might have been more used to tackling such types of questions in pairs and small groups and thus did not find the task too difficult and further, may have more accurately perceived the complexity of the task. Teachers in smaller classes usually spend more time in whole-group discussion as opposed to lecture, and on working with small groups than those in larger classes (Rice, 1999), and this pattern of instruction may have better-prepared the private-school students to work with a partner to answer a reasoning question. In which case, for answering the reasoning question, working with a partner did provide an advantage over working alone to answer the reasoning question.

### Conclusion

In this study, I have identified a number of factors that may be related to the success of peers engaged in unstructured collaborative learning with hypermedia and, for some, answering a reasoning question. The first factor that may be related to success is students’ perception that the task: (a) involves working with a peer to build a shared understanding; (b) is ill-structured; and (c) requires active learning processes in its execution. In this study, students’ perception of what the task entailed may have influenced their actions when performing the task.

The three examples of students collaborating offered in Chapter 4 provide evidence of several ways in which students mutually perceived the collaborative nature of the task—from merely sharing resources, to constructing a mutual understanding.

Similarly, some students answering the reasoning question treated the task as a search for

verbatim information, while others perceived answering the question as entailing more active processing of information, making connections and inferences, and constructing an answer based on what they had learned. The differences between public and private-school students that arose for learning gains and for the question response scores indicate that school experience may be related to how students perceive tasks (Blumenfeld, Mergendoller, & Swarthout, 1987). It is possible that the private-school students were exposed to a more critical-analytic approach in their academic life (owing to smaller class sizes), which may have affected how many of them approached the task.

The second factor related to success may be the use of high-quality regulatory processes. Students who engaged in these types of process (e.g., activating appropriate prior knowledge and making inferences about important content) were the students who had higher learning gains. Similarly, students need to maintain and regulate a high level of engagement in the task. More successful learners had low levels of off-task behavior in comparison to the high rate of off-task talk among those less successful. However, interest alone does not suffice and can often detract from learning important content, as seen in the high rate of interest in relatively insignificant details among less-successful students. Rather, the motivation needed to be sustained and coupled with an appropriate task perception and utilization of high-quality, active learning strategies.

The third factor that may be related to success in this type of context is some prior knowledge of the topic. While all students had low prior knowledge, some students struggled more than others to answer the reasoning question. Low prior knowledge appeared to affect their ability to find information in the hypermedia environment, as well as to monitor their answers to the question. However, these students also appeared to

view the task as a fact-finding activity rather than an inferential one. As such, prior knowledge may need to be coupled with an appropriate task perception to be fully utilized.

A fourth factor that may be related to student success in tasks such as this is time. This task was 30 minutes in duration, and it fit well within the constraints of classroom scheduling. However, it may not have left enough time for peers to discuss their plans and goals, share and build understanding, and to monitor their understanding. Further, putting a time limit on the task may have provided students with information about the nature of the task. It is possible that students interpreted such a relatively short task as one in which effort and active learning processes were not necessary.

In conclusion, there appear to be several individual (i.e., task perception, ability to regulate learning processes, and prior knowledge) and contextual (i.e., time of task) factors that may be related to how successful students are when they engaged in peer collaboration with an ill-structured task. Students' prior academic experiences may play a crucial role in determining the individual factors that students bring to a collaborative pair. The results of this study serve to highlight the potential importance of these factors, and in the next section, I address the implications and contributions of the results.

#### Implications and Contributions

Peer collaboration and questioning are oft-touted pedagogical methods that are frequently used in various instantiations in educational practice. The de facto assumption appears to be that peer collaboration fosters learning to a greater degree than more traditional independent learning. The results of this study call into question the broad assumption that these methods, as conceived in this study, are necessarily advantageous

for learning relative to their absence. As such, a number of implications for practice arise from the results of this study.

First, careful consideration should be given to using peer collaboration as a teaching technique. Although research has demonstrated that particular forms of scripted collaboration can help students (e.g., Slavin, 1996), anecdotal evidence suggests that short, unstructured peer collaboration is the more common type of peer collaboration implemented in classrooms. Based on the results of this study, this type of collaboration may not always be any more advantageous to academic learning than students working independently. The task in this study was open-ended, in the sense that nothing was scaffolded for the students, and their overall learning goal was very general (e.g., “learn all that you can about the circulatory system”).

If teachers wish to use unstructured collaboration, they should be aware that students have varied task perceptions and that they may be related to the processes students use when learning. Educators should be aware that students’ task perceptions and definitions appear to influence whether they approach the task using surface-level or deep approaches, and that these perceptions may mediate the success of peer learning and of questioning. The qualitative analysis of the coded transcripts revealed that several students interpreted (or chose to treat) the collaborative learning task as a sharing-of-resources task. Teachers employing collaboration in the classroom should make the process of collaboration apparent to students and encourage meaningful interaction between them. Classroom experience and support for active learning processes may help develop students’ appropriate perceptions of the task.

Teachers can be made aware of the particular processes utilized by the students who were more and less successful learners when working with a peer. Behaviors associated with more successful peer learners included active learning strategies such as summarizing and taking notes, and monitoring behaviors such as evaluating the content as they read and identifying feelings of knowing. Teachers can encourage and foster these behaviors in their students. In contrast, those who were less successful expressed more comments indicating they did not know or understand the information, and they engaged in far greater instances of off-task behavior. They also appeared to be distracted by relatively unimportant details as they read. If teachers are aware of these potential pitfalls of unstructured collaborative learning, they may be able to provide support and opportunities for students to practice these skills and possibly advance their epistemic stance.

Similarly, many students treated the question-answering task as a search for an answer only, with little attention to remembering the material they were reading. Teachers can model answering the question in a way that shows students how to build on their knowledge base at the same time. Educators should carefully consider how they use reasoning questions to help students learn, particularly with hypermedia and other non-linear sources of information. In an effort to address difficulties students have in computer-based learning environments, a number of researchers have advocated using conceptual scaffolds in the form of static questions (including reasoning questions) to help students learn (e.g., Hmelo & Day, 1998; Linn, 1996; Schank, Linn, & Clancy, 1993). Based on the results of this study, caution should be exercised before the wholehearted acceptance of such supports. In this study, the hypermedia environment

presented a challenge to many students, who spent more cognitive effort navigating and searching for the answer to the reasoning question than learning about it. Many students appeared to inaccurately perceive the aim of the question due possibly to a lack of experience or training answering such questions. Students with some prior knowledge of the topic and who have experience answering such questions may be best positioned to take advantage of the power of these questions. In this case, reasoning questions may help students activate their prior knowledge and set goals as they are learning.

This study contributes to the literature and research on collaborative learning and on questioning in several ways. It identified several factors that appear to relate to student learning. For example, students' task perceptions may relate to how they worked at the task. In this study, students' perceptions were often contrary to the intended goals of the instructor and researcher.. Likewise, this study identified the potential importance of prior knowledge in answering reasoning questions, particularly when using hypermedia. In this study, the students had relatively low prior knowledge about the circulatory system. As such, many students spent much of their limited time searching the environment for an answer to the question. Their low prior knowledge, coupled with an inaccurate task perception, may have meant that students did not always know what they did not know, where to look for the information they wanted, whether the information they were finding was adequate, and when their answer was complete.

The results of this study can help clarify particular factors that may be related to successful collaborative learning for conceptual understanding in science. In particular, this study found that students who learned the most did engage in particular processes related to performance and monitoring to a greater degree than those who learned the

least. Many of the strategic processes that high-gain students utilized to a greater extent are processes that can be scaffolded for students, such as summarizing and taking notes, until they are ready to engage in such processes independently (Alexander, 2005).

Students who were least successful learning collaboratively tended to have more trouble understanding the content and went off-task much more often.

This study also theoretically and methodologically contributes to research on individual and collaborative self-regulated learning from a social-cognitive perspective. In particular, this study investigated the triadic interaction between individual learners, an environmental context of peer learning with a hypermedia environment, and the learners' behavior. Few studies on peer collaboration and on learning with hypermedia have utilized a social-cognitive perspective, much less one with a focus on collaborative-regulated learning. As such, this study provides an initial foray into using social-cognitive theory to investigate these particular contexts. Methodologically, this study successfully utilized a collaborative self-regulatory coding scheme based on social-cognitive theory that worked to identify the regulatory processes students used as they learned. However, the results of this study also highlight that assessment of the quality of the students' regulatory behaviors is as important as identifying those processes. While the coding scheme used in this study was not able to describe the quality of students' verbalizations, qualitative analysis of the actual discourse between pairs began to reveal this. Further, the results of this study extend previous research on self and collaborative regulation of learning by identifying the potential role that task perception plays in determining students' regulatory processes and actions. This study builds on current theories of collaborative learning by using social-cognitive theory, which focuses on students'

regulation of their learning, rather than just the cognitive, metacognitive, or motivational aspects of collaborative learning in isolation.

Finally, this study contributes to the growing field of research on learning with multimedia and hypermedia. It provides additional support for the contention that students may have difficulty regulating their learning while using hypermedia and multimedia (Winters et al., 2008). Further, it extends the research conducted on the basic principles of multimedia and hypermedia learning conducted by Mayer and colleagues (Mayer, 2001, 2005) by using a more complex hypermedia environment. In complex environments, with complex and open-ended tasks, factors other than the cognitive ones considered by Mayer and colleagues (Mayer, 2001, 2005), such as students' motivations, become critical to students' success.

#### Limitations

This study has several limitations. Data were collected at three different schools (2 private and 1 public) over the course of six months. School-type effects were found, and in the discussion, I explored several potential explanations for this unanticipated effect. However, because of the study design—with private school data collected in late spring, and public school data collected in late fall, a cohort effect cannot be ruled out. Similarly, although I maintained a strict experimental protocol at each school, the settings differed. Random assignment of students to the experimental conditions within classrooms at each school rules out between-class differences, but the same cannot be said of the school-level effects.

Further, aspects of this study were quasi-experimental, with statistical groupings formed after the data had been collected. These statistical groups were determined by the sample, with the larger gain group constituting students whose learning gains were in the

upper 30% of the sample, and the smaller gain group constituting students whose learning gains were in the lower 30%. It is conceivable that a sample in which students made higher gains might engage in different regulatory processes.

To account for prior knowledge, I used a pretest-posttest design. Although this design provides needed information about students' prior knowledge, the presence of a pretest can alter students' actions in the task in several ways. First, the pretest may act as a prompt for prior knowledge once students start the task. However, that prior knowledge may not be evident on the actual pretest. In such a case, the pretest would underestimate a students' prior knowledge, and their gain scores would then be overestimated. Second, pretesting may well give students clues as to the content of the posttest, and will then drive their learning to the extent they wish to do well on the posttest. In other words, the pretest may guide learning. However, while this is a limitation, the alternative (no pretest) would have made estimates of any prior knowledge impossible.

Another limitation of this study is that the coding scheme does not adequately capture the quality of the processes it measures. To account for this, more in-depth qualitative analysis of three transcripts were included; however, it would have been too intensive to provide this level of analysis for every transcript. The exclusion of most of the transcripts from the dyad-level qualitative analysis is a limitation. Similarly, I could only investigate the processes involved with peers talking to each other. I was not able to capture the internal processes of these students, nor did I conduct concurrent think-alouds with the students working independently. This limits the ability to make true comparisons between peers' and individuals' learning processes.

Finally, other than prior knowledge, this study did not account for measurement of differences between students on factors, such as status, task perception, or epistemic beliefs that may have affected their participation in the task. Finally, the task the students engaged in for this study was short (i.e., only 30 minutes in duration). This must be taken into account when interpreting the results, as a longer learning activity would likely produce a different result.

#### Future Research

To build on the results of this investigation into peer collaboration and questioning, several avenues for future research exist. The results from this study provide a start to exploring ways of scaling up the process of conceptual understanding through peer collaboration. However, while this study identified a number of collaborative processes associated with conceptual understanding, future research should seek to build on these results and determine more conclusively the mechanisms behind successful peer collaboration.

The role of students' task perceptions, and how they determines the processes students use as they are learning is an important area for future research. Under the Pintrich (2000) framework, perception of the task occurs as an initial step in the forethought phase. The students' task definitions, based on their perceptions, likely has an effect on the goals they set, how they set about to achieve those goals, and their motivational mindset and behaviors. In short, this perception and definition may set the students' subsequent course of regulatory actions. As such, measurement of students' task perceptions is essential in future research on unstructured peer collaboration.

Related to this, the effect of school type or school environment is another contextual factor worth investigating, in light of the results of this study. The finding that

students' learning gains differed depending on their school type was unexpected and should be explored further in a more controlled manner. Was this difference due to a difference in school environment? Does a smaller class size provide students with greater opportunity to practice collaborative skills or answer reasoning questions? Does experience answering teacher-directed questions help foster higher-order self-questioning? How generalizable are these results to other schools and other students? The answers to these questions have important pedagogical and policy implications.

Can students be taught to engage in successful collaboration and reasoning question-answering? One approach to this inquiry could focus on whether providing scaffolding for the processes exhibited by the larger gain students help other students learn when working collaboratively. In this study, the association between learning and these processes was established, but whether these processes led to the observed learning gains should be empirically tested. The collaborative self-regulatory coding scheme provides a methodological tool that can be refined and extended to this end. Going forward, this research strand should focus on the quality as well as quantity of the learning processes students use while working collaboratively. As identified in this study, the quality of the behavior is vital to determining its efficacy.

Another area of potential future research is the role that contextual factors play in peer collaboration and questioning. This includes determining in what ways the hypermedia environment may have affected the way in which peers collaborated and the way in which students answered the reasoning question. Students spent much of their time searching the environment to find an answer to the question, so it would be worthwhile to determine whether a different type of source would have mitigated this

searching behavior. Further, in this study, students did not receive feedback about the correctness or completeness of their answers, so it was up to them to gauge when they thought they had completed the task of answering the question. The result was that many students did not write a thorough answer, likely because they were unaware that more answers existed due to a limited task perception. Would providing students with information about the correctness or completeness of their answer as they worked help them learn and re-evaluate their perception of the task? The impact that feedback about answers plays in student learning remains a viable path of research. This is particularly important given that questioning is often considered a type of scaffolding in hypermedia environments (Hannafin, Land, & Oliver, 1999), yet it has not received much research attention. Such research could determine the best ways to implement such supports in non-linear environments like hypermedia.

Another contextual factor that should be researched is the role the task itself plays in determining how successful students are at learning while collaborating. In this study, the students had 30 minutes for the task. Would extending the time of the task alter how students worked together and how they learned? Would a task that was more well-structured than the overall learning goal provided to students in this task affect how they worked together and how they learned? These questions arise from the results of this study, but remain unanswered.

The roles that individual factors play in peer learning and questioning also provide avenues for future research. A number of these factors were identified as potential influences during this study. For example, students' prior knowledge appeared to play a role in students' relative inability to answer the reasoning question adequately.

Future research should more systematically account for the ways in which prior knowledge affects how students answer reasoning questions. Similarly, students' reading level may affect how they learn when the task entails much reading to learn, as was the case in this study. Are there developmental differences in how students collaborate and learn in a reading-rich task? It is possible that the results presented in the present study might have been different with a more skilled sample of readers, such as those at the college level.

This study is an initial foray into applying social-cognitive theory to collaborative learning. While it provides some evidence that students do engage in collaborative regulation, to fully flesh out how well social-cognitive theory can explain peer collaboration, other factors need to be considered. In particular, self-efficacy is a key construct of social-cognitive theory (Bandura, 1986). The addition of measurement of students' self-efficacy for the task and for collaborating may reveal important differences between students that help explain varied success at the task.

From a methodological perspective, future research should focus on refinement of several of the measures and coding used in this study. In particular, the conceptual-knowledge measure could be refined. For the purposes of the present study, the total score was not parsed, as it was deemed to represent conceptual knowledge as a whole. Future research might involve investigating whether the score could be parsed, with question 6 (path diagram of the blood flow) constituting the measure of conceptual knowledge and questions 1-5 indicating declarative knowledge only. Counterbalancing these components would be necessary to avoid any fatigue effects. Last, potential

pretesting effects with this measure should be tested. This could be achieved by comparing two randomly-assigned groups—one with a pretest and the other without.

Another future methodological focus should be development of a coding scheme for the collaborative discourse that includes some identification of the quality of the behavior exhibited. The challenge this presents lies in the very contextual nature of this identification. Determination about whether a particular behavior is of high or low quality very much depends on what the student is doing and thinking at that particular moment, as well as what they have just done or thought prior to the behavior under inspection. Further, development of such a scheme requires a good understanding of the topic or domain the student is learning. Ideally, however, such a scheme would be relatively generalizable to a variety of topics and domains.

Together with the results of the current study, the strands of future research described here would not only contribute to the existing literature on peer collaboration and on questioning, but they have important implications for educators inclined to use peer collaborative or teacher-directed questioning methods.

## APPENDICES

## Appendix A: Conceptual-Knowledge Pretest and Posttest

**Pretest /(Posttest)****Participant ID:** \_\_\_\_\_**Date:****Condition:****Page 1 of 3**

---

Please read each question carefully and answer them to the best of your knowledge. The amount of space provided is NOT an indication of how much or how little you need to write. If you need more space, use the back of the sheet.

1. List the **function(s)** of the circulatory system:

---

---

---

---

---

---

---

---

---

---

---

2. List the **parts** of the circulatory system:

---

---

---

---

---



**Pretest / (Posttest)**

**Participant ID:** \_\_\_\_\_

**Date:**

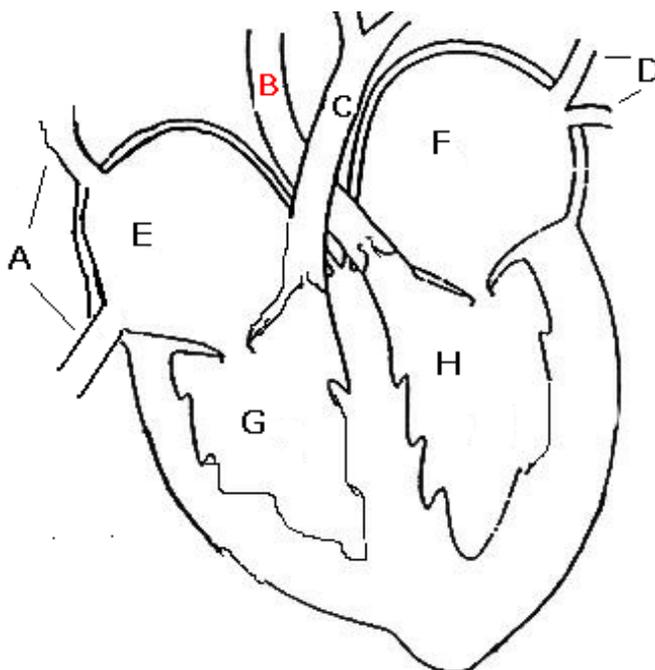
**Condition:**

**Page 3 of 3**

Please use the accompanying diagram to answer the following questions:

**5.** Write the name of each part of the heart next to the corresponding letter in the table.  
OR, you may write directly on the diagram if you prefer.

A	
B	
C	
D	
E	
F	
G	
H	



**6.** Starting from location **B**, describe the path of blood flow through the body part shown above. **Include any other major body parts NOT shown on this diagram, if necessary.**

You may use arrows to indicate directionality. For example: X → Y → Moon → Z → (etc.)



## Appendix C: Scoring Rubric for Pretest and Posttest Conceptual Knowledge

1. [Total possible = 10]
  - transports O<sub>2</sub> (2 pt.)
  - transports nutrients (2 pt.)
  - transport minerals, vitamins (1pt.)
  - body temp (1 pt)
  - hormones (1 pt)
  - heal cuts/clots blood (1 pt)
  - attack germs(1 pt) or works with immune system (.5)
  - carry away waste (1 pt)
  
2. [Total possible = 8]
  - Heart (1)
  - List parts (at least 2) of heart (1)
  - Blood (1)
  - List parts of blood (at least 2) (1)
  - Blood vessels (1)
  - List particular blood vessels (at least 2) (1)
  - Lungs (1)
  - Valves in heart or veins (1)
  
3. [Total possible = 10]

<b>Part(s)of Blood</b>	<b>Function(s) of the Part(s)</b>
White Blood Cells (1)	Fight Disease/ germs (1)
Hemoglobin (.5) or Red Blood Cells (1)	Carry Oxygen (1) and Carbon Dioxide (1)
Plasma (1)	Carry Nutirents (1) water, hormones, minerals, waste (any one or more mentioned) (1)
Platelets (1)	Heal cuts (1) or (helps blood clot)

4. [Total possible = 7]

<b>Part(s) of the Blood Vessel System</b>	<b>Function(s) of the Part(s)</b>
Arteries (1)	Carry blood away from the heart (1)
Veins (1)	Carry blood towards the heart (1)

Capillaries (1)	Connect arteries to veins (1); Place where nutrients, gasses are exchanged (1)
Valves (.5)	In veins, keep blood from flowing backwards (.5)

5. [Total possible = 8]

A	Vena Cavae (1) [.5 if only superior or inferior or if pulmonary artery]
B	Aorta (1) [.5 if P. artery]
C	Pulmonary Artery (1) [.5 if p. vein OR aorta]
D	Pulmonary Vein (1) [.5 if p. artery OR Inf. Vena cava]
E	Right Atrium (1) [.5 if L. atrium OR R. ventricle]
F	Left Atrium (1) [.5 if R. atrium OR L. ventricle]
G	Right Ventricle (1) [.5 if L. ventricle OR R. atrium]
H	Left Ventricle (1) [.5 if R. ventricle OR L. atrium]

6. B → Body → A → E → G → C → Lungs → D → F → H → B  
           2      2      1   1   1   3          3   1   1   1

(total of 16 points possible)

(reversed - right and left sides confused (NOT reversed directionality of flow) = - 4 points )

(If A/D and B/C are reversed = -4 BUT MUST include LUNGS)

(3 points if lungs are written in any part of path)

*Total Measure: 60 points possible*

## Appendix D: Coding Scheme for Peer Discourse

*Classes, Descriptions and Examples of the Variables Used to Code Learners' Collaborative Regulatory Behavior (adapted from Azevedo, Cromley, Winters, Moos, & Greene, 2005; Winters & Azevedo, 2005)*

Phase Variable	Description <sup>1</sup>	Examples (from Azevedo, Cromley, Winters, Moos, & Greene, 2005; Winters & Azevedo, 2005)
<i>Planning (Forethought)</i>		
Prior Knowledge Activation (PKA)	Searching memory for relevant prior knowledge from a previous class or experience.  Verbalizes content from prior knowledge	"I know that people with anemia get tired more often than people without it"
Recycle Goal (RG)	Restating the goal (TSG or G) in working memory	
SubGoal (SG)	Stating a goal for learning that is different from the overall TSGs.	"Let's find out more about the heart"  "I want to know what blood type has to do with this"
Teacher-Set Goals (TSG)	Verbalization of a question or set of directions provided by the teacher (such as those on a worksheet).	<i>"If you were a blood cell located for the moment in a person's big toe, it would take you longer to travel from the toe up to the left side of the person's heart than for you to go from the left side of the heart back down to the toe. Why? Name as many reasons as possible"</i>  <i>"Your task is to learn all you can about the circulatory system in 30 minutes, be sure to learn about the parts and their purpose, how they work both individually and together to support the human body"</i>
Time and Effort Planning (TEP)	Attempts to intentionally control behavior related to time or effort.	"Let's just finish this one."

<sup>1</sup> All codes refer to what was recorded in the verbal protocols (i.e., what students read, saw, or heard)

<i>Performance (Control)</i>		
Coordinating Informational Sources (COIS)	Coordinating multiple representations, e.g., drawing and notes.	"I'm going to put that [text] with the diagram"
Inference (INF)	Student makes an inference based on text reading	"So the increased pressure in the arteries must make the blood move faster"
Elaboration/explanation (KE)	Student provides <b>explanation</b> or comment using PK about something they read, beyond what is provided in the text	"It says it takes 30 seconds for the blood to travel around the whole body – <b>that is fast</b> "
Memorize (MEM)	Learner tries to memorize text, diagram, etc. or use a mnemonic device	"I'm going to try to memorize this picture" "Arteries – away from heart – ok, 'a' away"
Partner Questioning for Procedure (PQP)	Student asks partner question pertaining to the procedure of the task (not a conceptual question)	"That's what the question is asking?" "How do I get back to that diagram?"
Read out Loud (RL)	Learner reads out loud for his and partner's benefit	
Read Notes (RN)	Reviewing learner's notes.	"I am reading over my notes."
Seeks Consensus (SC)	Student seeks partner's agreement on conceptual information or on procedural decision. (Student asking is sure of own convictions)	"Can we go back?" "Hey, you wanna do this one?"
Search (SEARCH)	Searching the hypermedia environment after specifying a specific goal OR using the search feature	Learner types in "blood circulation" in the search feature
Summarize (SUM)	Student summarizes text or diagram.	"So, the three big parts of the circulatory system are the heart, blood and blood vessels"
Take Notes (TN)	Taking notes on material being learned [often evidenced in physical notes rather than a verbalization]	"I'm going to write that under heart"

<i>Monitoring (Reflection)</i>		
Content Evaluation (CE +, -)	Stating that any just-seen text, diagram or video is relevant (+) or irrelevant (-) to the learning task	<p>“That section was good for finding out about the heart” (CE+)</p> <p>“That section didn’t help us” (CE -)</p>
Expectation of Adequacy of Content	Expecting that a certain type of representation (usually text) will prove adequate given the current goal.	“This section/picture will help me answer the question”
Feeling of Knowing (FOK+, -)	<p>Learner makes a statement indicating he or she is aware of having learned something in the past and having some understanding of it (it is familiar +) or not (it is unfamiliar -).</p> <p>Different from PKA, which is initiated by student and includes content - this is more of an overall statement about whether the information is familiar or not.</p>	<p>“This tells us what we already know” (FOK +)</p> <p>“I knew that already” [But PM+ if referring to answer put on pretest] (FOK +)</p> <p>“We learned about that last year” (FOK +)</p> <p>“I don’t remember learning this last year” (FOK -)</p>
Judgment of Learning (JOL+, -)	Learner becomes aware that they do (+) or do not (-) know or understand something about what they are learning while engaged in the task	<p>“Uh, I don’t know” (JOL -)</p> <p>“I don’t understand this” (JOL-)</p> <p>“We know that now” (JOL +)</p>
Monitor Progress Toward Goals (MPTG)	Assessing whether previously-set goal has been met	<p>“Okay, I think we are finished answering the question”</p> <p>“We already have all the information we need”</p>
Partner Questioning for Understanding (PQU)	Learner directs conceptual question towards partner	<p>“But why does the blood go to the lungs?”</p> <p>“Doesn’t the blood go right out to the body after going to the lungs?”</p>

Pretest Monitoring (PM +, -)	Learner assess whether they got and answer right (+) or wrong (-) based on what they have just learned	“I got that one right on the test” (PM +) “Oh, I didn’t put that one right on the pretest” (PM -)
Seeks Affirmation (SA)	Student seeks agreement from partner about a conceptual idea that he/she is unsure about – usually preceded by PQP or PQU.	“Uh, for the last one we say yes, <b>right</b> ? “So, the original... <b>right</b> ?”
Self Correct (SEC)	Learner realizes they made a mistake or misjudgment (can pertain to hypotheses or result being “wrong”).	“I guess I was wrong” “We could just say right – oh, wait, it does say left!”
Strategy Monitoring (SM)	Learner makes a statement about the adequacy of a particular learning strategy, such as taking notes.	“I need to take notes because it helps me remember stuff”
Time Monitoring (TM)	Participant indicates that a certain amount of time is left in the learning task	“We only have 5 minutes left.”
<i>Motivation</i>		
Off-task (OT)	Learner exhibits behavior that is clearly not within the learning task, including talking off-topic with collaborative partner and/or peer.	“She dresses nice.” “I am so craving some chicken right now.”
Positive feedback (PF)	Partner gives encouragement or affirmation for an idea, choice, or question (PF)	“Yeah.” “Right.”
Negative feedback (NF)	Partner discourages or disagrees (NF) [usually in response to PQU or PQP]	“ No, ----“ “I don’t think so”
Positive interest (Int +)	Learner expresses excitement or interest in task or aspect of task.	“Oh my God!” “Wow!” “Eww – gross”
Negative interest (Int -)	Learner expresses boredom or lack of interest in task or aspect of task.	“I really like this project – or NOT.”

## Appendix E: Scoring Rubric for Answers to Reasoning Question

0	No answer OR none of the below.
1	Blood moves more slowly going up than down.
2	-Gravity slows the blood down going up and helps it going down. OR -It is a longer route going up than down OR -There is more pressure pushing the blood on the way down (No further explanation/reasoning of why)
3	Two or more of category 2, together.
4	-There is a longer route going up because the blood must travel through the pulmonary system (through the right side of the heart, then out to the lungs and back to the left side of the heart) first. When the blood leaves the left side, it goes directly through the systemic system to the toe. [can also include any from category 1-2]
5	There is more pressure pushing the blood on the way down because it is closer to the left ventricle, which gives a strong push when the blood leaves the heart. By the time the blood is returning to the heart, the pressure has dissipated. OR If a person is standing, muscles aren't helping to push the blood back up to the heart, so the blood will move more slowly through the veins.
6	Any 2 from category 4 and/or 5 (can include gravity explanation)
7	All 3 from category 4 and 5 (can include gravity explanation)
8	7 + gravity explanation

## REFERENCES

- Alexander, P. A. (1995). Superimposing a situation-specific and domain-specific perspective on an account of self-regulated learning [Special issue] *Educational Psychologist*, 30, 189-193.
- Alexander, P. A. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher*, 32(8), 10-14.
- Alexander, P. A. (2005). The path to competence: A lifespan developmental perspective on reading. *Journal of Literacy Research*, 37(4), 413-436.
- Alexander, P. A., Kulikowich, J. M., & Jetton, T. L. (1994). The role of subject-matter knowledge and interest in the processing of linear and non-linear texts. *Review of Educational Research*, 64(2), 201-252.
- Alexander, P. A., & Jetton, T. L. (1996). The role of importance and interest in the processing of text. *Educational Psychology Review*, 8(1), 89-121.
- Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research*, 61(3), 315-343.
- Alexopoulos, E., & Driver, R. (1996). Small-group discussion in physics: Peer interaction modes in pairs and fours. *Journal of Research in Science Teaching*, 33(10), 1099-1114.
- American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.
- American Association for the Advancement of Science. (1997). *Blueprints: Project 2061*. Washington, DC: AAAS.

- Anderson, R. C., & Biddle, W. B. (1975). On asking people questions about what they are reading. In G. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9, pp. 89-132). New York: Academic Press.
- Andre, T. (1979). Does answering higher-level questions while reading facilitate productive learning? *Review of Educational Research*, 49(2), 280-318.
- Azevedo, R. (2005). Computers as metacognitive tools for enhancing learning. *Educational Psychologist*, 40(4), 199-209.
- Azevedo, R., & Cromley, J. G. (2004). The role of self-regulated learning in fostering students' understanding of complex systems with hypermedia. *Journal of Educational Psychology*, 96, 523-535.
- Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology*, 29(3), 344-370.
- Azevedo, R., Cromley, J. G., Winters, F. I., Moos, D. C., & Greene, J. A. (2005). Adaptive human scaffolding facilitates adolescents' self-regulated learning with hypermedia. *Instructional Science*, 33, 381-412.
- Azevedo, R., Winters, F. I., & Moos, D. C. (2004). Can students collaboratively use hypermedia to learn about science? The dynamics of self- and other-regulatory processes in an ecology classroom. *Journal of Educational Computing Research*, 31(3), 215-245.
- Baddeley, A. & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 8, pp.17-90). Orlando, FL: Academic Press.

- Bandura, A. (1978). The self system in reciprocal determinism. *American Psychologist*, 33, 344-358.
- Bandura, A. (1986). *Social foundations of thought and action: A social-cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bianchini, J. A. (1997). Where knowledge construction, equity, and context intersect: Student learning of science in small groups. *Journal of Research in Science Teaching*, 34(10), 1039-1065.
- Blumenfeld, P. C., Mergandoller, J. R., & Swarthout, D. W. (1987). Task as a heuristic for understanding student learning and motivation. *Journal of Curriculum Studies*, 19(2), 135-148.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.) (1999). *How people learn: Brain, mind, experience, and school*. Washington DC: National Academy Press.
- Brown, A. L., Bransford, J. D., Ferrara, R., & Campione, J.C. (1983) Learning, remembering, and understanding. In J. H. Flavell & E. M. Markham (Eds.), *Handbook of child psychology* (Vol. 3, pp. 77-166). New York: Wiley.
- Byrnes, J. P. (2001). *Cognitive development and learning in instructional contexts* (2nd edition). Needham, MA: Allyn & Bacon.
- Calisir, F., & Gurel, Z. (2003). Influence of text structure and prior knowledge of the learner on reading comprehension, browsing, and perceived control. *Computers in Human Behavior*, 19(2), 135-145.
- Callender, A. A., & McDaniel, M. A. (2007). The benefits of embedded question adjuncts for low and high structure builders. *Journal of Educational Psychology*, 32, 339-348.

- Carter, G., & Jones, M. G. (1994). Relationship between ability-paired interactions and the development of fifth graders' concepts of balance. *Journal of Research in Science Teaching, 13*(8), 847-856.
- Chan, C. K. K. (2001). Peer collaboration and discourse patterns in learning from incompatible information. *Instructional Science, 29*, 443-479.
- Chi, M. T. H., Bassok, M., Lewis, M., 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 Leeuw, N., Chui, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439-477.
- Chi, M. T. H., & Ohlsson, S. (2005). Complex declarative learning. In K. J. Holyoak & R. G. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 371-399). Cambridge, England: Cambridge University Press.
- Chin, C., & Brown, D. E. (2000). Learning in science: A comparison of deep and surface approaches. *Journal of Research in Science Teaching, 37*(2), 109-138.
- Chin, C., Brown, D. E., & Bruce, B. C. (2002). Student-generated questions: a meaningful aspect of learning in science. *International Journal of Science Education, 24*(5), 521-549.
- Chinn, C. A., O'Donnell, A. M., & Jinks, T. S. (2000). The structure of discourse in collaborative learning. *Journal of Experimental Education, 69*, 77-97.

- Cohen, E. G. (1984). Talking and working together: Status interaction and learning. In P. Peterson, L. C. Wilkinson, & M. Hallinan (Eds.), *Instructional groups in the classroom: Organization and processes* (pp.171-188). Orlando: Academic Press.
- Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64(1), 1-35.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Collins, A., & Ferguson, W. (1993). Epistemic forms and epistemic games: Structures and strategies to guide inquiry. *Educational Psychologist*, 28, 25-42.
- Costa, J., Caldera, H., Gallastegui, J. R., & Otero, J. (2000). *Journal of Research in Science Teaching*, 37(6), 602-614.
- Craig, S. D., Sullins, Witherspoon, A. & Gholson, B. (2006) The deep-level reasoning questions effect: The role of dialogue and deep-level reasoning questions during vicarious learning. *Cognition and Instruction*.24(4), 565-591.
- De Jong, T., & Van Jooligen, W.R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, 68(2), 179-201.
- De Lisi, R. (2002). From marbles to instant messenger: The implications of Piaget's ideas about peer cooperation for educational practice. *Theory Into Practice*, 41, 5-12.
- DeLeeuw, N. & Chi, M. (2003). Self-explanation: Enriching a situation model or repairing a domain model? In G.Sinatra & P.Pintrich (Eds.), *Intentional conceptual change* (pp. 55–78). Mahwah, NJ: Lawrence Erlbaum Associates.

- Derry, S., & Lajoie, S. (1993). *Computers as cognitive tools*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- De Vries, E., & de Jong, T. (1997). Using information systems while performing complex tasks: an example from architectural design. *International Journal of Human-Computer Studies*, 46, 31-54.
- Dillenbourg, P. (1999). Introduction: What do you mean by “collaborative learning”? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1-19). Oxford, England: Elsevier.
- Dillon, A., & Gabbard, R. (1998). Hypermedia as an educational technology: A review of the quantitative research literature on learner comprehension, control, and style. *Review of Educational Research*, 68(3), 322-349.
- Dillon, A., & Jobst, J. (2005). Multimedia learning with hypermedia. In R. Mayer (Ed.). *Cambridge handbook of multimedia learning* (pp. 569-588). New York, NY: Cambridge University Press.
- Driver, R. Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3), 287-312.
- Dunbar, K., & Fugelsang, J. (2005). Scientific thinking and reasoning, In Holyoak, K. J. & Morrison, R.G. (Eds.). *Cambridge handbook of thinking and reasoning* (pp. 705-725). Cambridge, England: Cambridge University Press.
- Engelkamp, J. (1998). *Memory for actions*. Hove, England: Taylor & Francis Psychology Press.
- Englehart, J. M. (2007). Discourse in a small class: The ‘diverge-converge pattern’ and ‘relaxed freedom’. *Education 3-13*, 35(1) 83-97.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*.

Cambridge, MA: The MIT Press.

Fantuzzo, J. W., King, J. A., & Heller, L. R. (1992). Effects of reciprocal peer tutoring on mathematics and school adjustment: A component analysis. *Journal of Educational Psychology, 84*(3), 331-339.

Faulkner, D., Joiner, R., Littleton, K., Miell, D., & Thompson, L. (2000). The mediating effect of task presentation on collaboration and children's acquisition of scientific reasoning. *European Journal of Psychology of Education, 15*(4), 417-430.

Finn, J. D., Pannozzo, G. M., & Achilles, C. M. (2003). The "whys" of class size: Student behavior in small classes. *Review of Educational Research, 73*(3), 321-368.

Forman, E. A. & Cazden, C B. (1994). Exploring Vygotskian perspectives in education: The cognitive value of peer interaction. In R. B. Ruddell, RB, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp.155-178). Newark, DE: International Reading Association.

Frederiksen, N. (1984). Implications of cognitive theory for instruction in problem solving. *Review of Educational Research, 54* (3). 363-407.

Garner, R., Gillingham, M. G., & White, C. S. (1989). Effects of "seductive details" on macroprocessing and microprocessing in adults and children. *Cognition and Instruction, 6*(1), 41-57.

Graesser, A. (1993). *Questioning mechanisms during complex learning*. Memphis State University, Memphis, TN (ERIC Document Reproduction Service No. (ED 350 306).

- Graesser, A., Baggett, W., & Williams, K. (1996). Question-driven explanatory reasoning. *Applied Cognitive Psychology, 10*, S17-S31.
- Graesser, A. C., & Franklin, S. P. (1990) QUEST: A cognitive model of question answering. *Discourse Processes, 13*, 279-303.
- Graesser, A., McNamara, D., & VanLehn, K. (2005). Scaffolding deep comprehension strategies Through Point & Query, AutoTutor, and iSTART. *Educational Psychologist, 40*(4), 225-234.
- Graesser, A., & Person, N. (1994). Question asking during tutoring. *American Educational Research Journal, 31*(1), 104-37.
- Graesser, A., Person, N., & Huber, J. D. (1992). Mechanisms that generate questions. In T. E Lauer, E. Peacock, & A. C. Graesser (Eds.), *Question and information systems* (pp. 167-187). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Graesser, A., Person, N., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology, 9*, 495-522.
- Gredler, M. E. (2004). Games and simulations and their relationships to learning. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2nd ed; pp. 577-591). Mahwah, NJ: Lawrence Erlbaum Associates.
- Grigg, W.S., Lauko, M.A., & Brockway, D.M. (2006). *The Nation's Report Card: Science 2005* (NCES 2006-466). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Grossen, M. (1994). Theoretical and methodological consequences of a change in the unit of analysis for the study of peer interactions in a problem solving situation. *European Journal of Psychology of Education, 9*(1), 159-173.

- Hacker, D. J., Tenent, A. (2002). Implementing reciprocal teaching in the classroom: Overcoming obstacles and making modifications. *Journal of Educational Psychology, 94*, 699-718.
- Hamaker, C. (1986). The effects of adjunct questions on prose learning. *Review of Educational Research, 56*(2), 212-242.
- Hannafin, M., Land, S., & Oliver, K. (1999). Open learning environments: Foundations, methods, and models. In C. Reigeluth (Ed.), *Instructional Design Theories and Models* (pp. 115-140). Mahwah, NJ: Lawrence Erlbaum Associates.
- Harp, S. F., & Mayer, R. E. (1997). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90*(3), 414-434.
- Hegarty, M. (2005). Multimedia learning about physical systems. In R. Mayer (Ed.). *Cambridge handbook of multimedia learning* (pp. 447-466). Cambridge, England: Cambridge University Press.
- Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction, 21*(4), 325-360.
- Hegarty, M., Quilici, J., Narayanan, N. H., Holmquist, S., & Moreno, R. (1999). Multimedia instruction: Lessons from evaluation of a theory-based design. *Journal of Educational Multimedia and Hypermedia, 8*(2), 119-150.
- Herrenkohl, L. R., & Guerra, M. R. (1998). Participant structures, scientific discourse, and student engagement in fourth grade. *Cognition and Instruction, 16*(4), 431-473.

- Hmelo, C., & Day, R. (1999). Contextualized questioning to scaffold learning from simulations. *Computers and Education, 32*(2), 151-164.
- Hmelo, C., Holton, D. L., & Kolodner, J. L. (2000). Designing to learn about complex systems. *Journal of the Learning Sciences, 9*(3), 247-298.
- Hmelo-Silver, C., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors and functions. *Cognitive Science, 28*(1), 127-138.
- Hogan, K. (1999). Thinking aloud together: A test of an intervention to foster students' collaborative scientific reasoning. *Journal of Research in Science Teaching, 36*(10), 1085-1109.
- Hogan, K., Nastasi, B. K., & Pressely, M. (1999). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction, 17* (4), 379-432.
- Hogan, D. M., & Tudge, J. R. H. (1999). Implications of Vygotsky's theory for peer learning. In A. M. O'Donnell and A. King (Eds.) *Cognitive perspectives on peer learning* (pp. 39-65). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Holliday, W. G. (1981). Selective attentional effects of textbook study questions on student learning in science. *Journal of Research in Science Teaching, 18*(4), 283-289.
- Holliday, W. G. (1983). Overprompting science students using adjunct study questions. *Journal of Research in Science Teaching, 20*(3), 195-201.
- Holliday, W. G., & Benson, G. (1991). Enhancing learning using questions adjunct to science charts. *Journal of Research in Science Teaching, 28*(1), 97-108.

- Holliday, W. G., & Mcguire, B. (1992). How can comprehension adjunct questions focus students' attention and enhance concept learning of a computer-animated science lesson? *Journal of Research in Science Teaching*, 29(1), 3-16.
- Holliday, W. G., Whittaker, H. G., & Loose, K. D. (1984). Differential effects of verbal aptitude and study questions on comprehension of science concepts. *Journal of Research in Science Teaching*, 21(2), 143-150.
- Jacobson, M. J. (1996). Learning with hypertext learning environments: Theory, design, and research. *Journal of Educational Multimedia and Hypermedia*, 5(3-4), 239-281.
- Jacobson, M. J., & Kozma, R. B. (2000). *Innovations in science and mathematics education: advanced designs for technologies of learning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning sciences. *Journal of the Learning Sciences*, 15(1), 11-34.
- Jeong, H., & Chi, M. T. H. (2007). Knowledge convergence and collaborative learning. *Instructional Science*, 35, 287-315.
- Johnson, D. W., & Johnson, R. T. (1991). *Learning together and alone: Cooperative, competitive, and individualistic learning*. Englewood Cliffs, NJ: Prentice Hall.
- Jonassen, D. H., & Reeves, T. C. (1996). Learning with technology: Using computers as cognitive tools. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 693-719). New York: Macmillan.

- Kaartinen, S., & Kumpulainen, K. (2002). Collaborative inquiry and the construction of explanations in the learning of science. *Learning and Instruction, 12*, 189-212.
- King, A. (1989). Effects of self-questioning training on college students' comprehension of lectures. *Contemporary Educational Psychology, 14*, 366-381.
- King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31*(2), 338-368.
- King, A., & Rosenshine, B. (1993). Effects of guided cooperative questioning on children's knowledge construction. *Journal of Experimental Education, 61*(2), 127-148.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, England: Cambridge University Press.
- Kneser, C., & Ploetzner, R. (2001). Collaboration on the basis of complementary domain knowledge: Observed dialogue structures and their relation to learning success. *Learning and Instruction, 11*, 53-83.
- Kozma, J. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction, 13*(2), 205-226.
- Kumpulainen, K., Salovaara, H., & Mutanen, M. (2001). The nature of students' sociocognitive activity in handling and processing multimedia-based science material in a small group learning task. *Instructional Science, 29*, 481-515.

- Lajoie, S.P. (2000). *Computers as cognitive tools ii: no more walls: Theory change, paradigm shifts and their influence on the use of computers for instructional purposes*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Lajoie, S.P., & Azevedo, R. (2006). Teaching and learning in technology-rich environments. In P. Alexander & P. Winne (Eds.), *Handbook of Educational Psychology* (2nd ed; pp.803-824.). Mahwah , NJ : Lawrence Erlbaum Associates.
- Lawless, K. A., & Kulikowich, J. M. (1996). Understanding hypertext navigation through cluster analysis. *Journal of Educational Computing Research*, 14(4), 385-399.
- Lemke, M., Calsyn, C., Lippman, L., Jocelyn, L., Kastberg, D., Liu, Y., et al. (2001). *Outcomes of learning: Results from the 2000 Program for International Student Assessment of 15-year-olds in reading, mathematics, and science literacy* (NCES 2002–115). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U. S. Government Printing Office.
- Lemke, M., Sen., A., Pahlke, E., Partelow, L., Mille, D., Williams, T., Kastbey, D., & Jocelyn, L. (2004). *International outcomes of learning in mathematics literacy and problem solving: PISA 2003 results from the US perspective* (NCES 2005-003). U.S. Department of Education, National Center for Education Statistics. Washington DC: U.S. Government Printing Office.
- Leonard, W., & Lowery, L. F. (1984). The effects of question types in textual reading upon retention of biology concepts. *Journal of Research in Science Teaching*, 21(4), 377-384.
- Linn, M. C. (1996). Cognition and distance learning. *Journal of the American Society for Information Science*, 47(11), 826-842.

- Lou, Y., Abrami, P. C., & Spence, J. C. (2000). Effects of within-class grouping on student achievement: An exploratory model. *Journal of Educational Research, 94*(2), 101-112.
- Lumpe, A. T., & Staver, J. R. (1995). Peer collaboration and concept development: Learning about photosynthesis. *Journal of Research in Science Teaching, 32*(1), 71-98.
- Martin, V.L., & Pressley, M. (1991). Elaborative-interrogation effects depend on the nature of the question. *Journal of Educational Psychology, 83*, 113-119.
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning: Outcome and process. *British Journal of Educational Psychology, 46*, 4-11.
- Mayer, R. E. (1997). Multimedia learning: are we asking the right questions? *Educational Psychologist, 32*(1), 1-19.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.
- 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*(2), 390-397.
- Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied, 11*(4), 256-265.

- Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: when presenting more material results in less understanding. *Journal of Educational Psychology, 93*(1), 187-198.
- 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*(1), 171-185.
- Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multi-media learning: evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*(2), 312-320.
- Mayer, R.E., Moreno, R., Boire, M., & Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology, 91*(4), 638-643.
- McCaslin, M., & Hickey, D. T., (2001). Self-regulated learning and academic achievement: A Vygotskian view. In B. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp.227-252). Mahwah, NJ: Lawrence Erlbaum Associates.
- Mislevy, R.J., Almond, R.G., & Lukas, J.F. (2004). A brief introduction to evidence-centered design. *CSE Technical Report 632*. Los Angeles: The National Center for Research on Evaluation, Standards, Student Testing (CRESST), Center for Studies in Education, UCLA.
- Mislevy, R.J., & Risconscente, M. (2006). Evidence-centered assessment design: Layers, concepts, and terminology. In S. Downing & T. Haladyna (Eds.), *Handbook of Test Development* (pp. 61-90). Mahwah, NJ: Erlbaum.

- Miyake, N., & Norman, D. A. (1979). To ask a question, one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior*, 18(3), 357-364.
- Moos, D.C., & Azevedo, R. (2006). The role of goal structure in undergraduates' use of self-regulatory processes in two hypermedia learning tasks. *Journal of Educational Multimedia and Hypermedia*, 15(1), 49-86.
- Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: the role of modality and contiguity. *Journal of Educational Psychology*, 93(1), 187-198.
- Moreno, R., & Mayer, R. E. (2002a). Verbal redundancy in multimedia learning: when reading helps listening. *Journal of Educational Psychology*, 94(1), 156-163.
- Moreno, R. & Mayer, R. E. (2002b). Learning science in virtual reality environments: role of methods and media. *Journal of Educational Psychology*, 94(3), 598-610.
- Moreno, R. & Mayer, R. E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology*, 97(1), 117-128.
- National Research Council (1996). *Science education standards*. Washington, DC: National Academy Press.
- National Center for Education Statistics (1998). Long-term trends in science performance. *NAEP Facts*, 3 (3), 1-4.
- O'Donnell, A. M. (2006). The role of peers and group learning. In P. Alexander & P. Winne (Eds.) *Handbook of educational psychology* (2<sup>nd</sup> ed; pp. 781-802). Mahwah, NJ: Lawrence Erlbaum Associates.
- O' Donnell, A. M., & Dansereau, D. F. (1992). Scripted cooperation in dyads: A method for analyzing and enhancing academic learning and performance. In R. Hertz-

- Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp.120-141). Cambridge, England: Cambridge University Press.
- O'Donnell, A. M., & O' Kelly, J. (1994). Learning from peers: Beyond the rhetoric of positive results. *Educational Psychology Review*, 6, 327.
- O' Donnell, A. M., & King, A. (1999). *Cognitive perspectives on peer learning*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Okada, T., & Simon, H. A. (1997). Collaborative discovery in a scientific domain. *Cognitive Science*, 21(2), 109-146.
- Osman, M.E., & Hannafin, M.J. (1994). Effects of advance questioning and prior knowledge on science learning. *Journal of Educational Research*, 88(1), 5-13.
- Ozgunor, S., & Guthrie, J. T. (2004). Interactions among elaborative interrogation, knowledge, and interest in the process of constructing knowledge from text. *Journal of Educational Psychology*, 96(3), 437-443.
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117-175.
- Palinscar, A. S., & Herrenkohl, L. R. (2002). Designing collaborative learning contexts. *Theory Into Practice*, 41, 26-32.
- Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, 76, 241-263.
- Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451-502). San Diego: Academic Press.

- Potelle, H., & Rouet, J. F. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. *International Journal of Human-Computer Studies*, 58(3), 327-345.
- Pressley, M., McDaniel, M. A., Turnure, J. E., Wood, E., & Ahmad, M. (1987). Generation and precision of elaboration: Effect on intentional and incidental learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(2), 291-300.
- Pressley, M., Symons, S., McDaniel, M. A., Snyder, B. L., & Turner, J. E. (1988). Elaborative interrogation facilities in the acquisition of confusing facts. *Journal of Educational Psychology*, 80, 301-342.
- Pressley, M., Tanenbaum, R., & McDaniel, M. A. (1990). What happens when university students try to answer prequestions that accompany textbook material? *Contemporary Educational Psychology*, 15(1), 27-35.
- Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist*, 27, 91-110.
- Recker, M. M., & Pirolli, P. (1995). Modeling individual differences in students' learning strategies. *Journal of the Learning Sciences*, 4(1), 1-38.
- Reed, S. K. (2006). Cognitive architectures for multimedia learning. *Educational Psychologist*, 41(2), 87-98.
- Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences*, 13(3), 273-304.
- Reynolds, R. E. & Anderson, R. C. (1982). Influence of questions on the allocation of

- attention during reading. *Journal of Educational Psychology*, 74(5), 623-632.
- Rice, J. K. (1999). The impact of class size on instructional strategies and the use of time in high school mathematics and science courses. *Educational Evaluation and Policy Analysis*, 21(2), 215-229.
- Riley, J. P. (1986). The effects of teachers' wait-time and knowledge comprehension questioning on science achievement. *Journal of Research in Science Teaching*, 23(4), 335-342.
- Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed). *Computer supported collaborative learning* (pp. 69-97). Germany: Springer-Verlag Berlin Heidelberg, NATO Scientific Affairs Division.
- Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research*, 64(4), 479-530.
- Rosenshine, B., & Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, 66(2), 181-221.
- Roy, M. & Chi, M. T. H. (2005). The self-explanation principle in multimedia learning. In R. Mayer (Ed.). *Cambridge Handbook of Multimedia Learning* (pp. 271-286). Cambridge, England: Cambridge University Press.
- Ryle, G. (1949). *The concept of mind*. Chicago: The University of Chicago Press.
- Sabelli, N. (2006). Complexity, technology, science, and education. *Journal of the Learning Sciences*, 15(1), 5-9.

- Saleh, M., Lazonder, A. W., & De Jong, T. (2005). Effects of within-class ability grouping on social interaction, achievement, and motivation. *Instructional Science*, 33, 105-119.
- Scardamalia, M., Bereiter, C. (1992). Text-based and knowledge-based questioning by children. *Cognition and Instruction*, 9(3), 177-199.
- Scardamalia, M., Bereiter, C., & Lamon, M. (1994). The CSILE project: Trying to bring the classroom into the world 3. In K McGilley (Ed.), *Classroom lessons, integrating cognitive theory and classroom practice* (pp.201-228). Cambridge, MA: MIT Press.
- Schank, P. K., Linn, M. C., & Clancy, M. L. (1993). Supporting Pascal programming with an on-line template library and case studies. *International Journal of Man-Machine Studies*, 38(6), 1031-1048.
- Schraw, G. (2006). Knowledge: Structure and processes. In Alexander, P. A. & Winne, P. (Eds), *Handbook of Educational Psychology: 2<sup>nd</sup> Edition*, (pp.245-264). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schunk, D. (1987). Peer models and children's behavioral change. *Review of Educational Research*, 57(2), 149-174.
- Schunk, D. (2001). Social cognitive theory of self-regulated learning. In & B. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 125-152). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schunk, D. (2005). Commentary on self-regulation in school contexts. *Learning and Instruction*, 15(2), 173-177.

- Schunk, D., & Swartz, C. W. (1993). Goals and progress feedback: Effects on self-efficacy and writing instruction. *Contemporary Educational Psychology, 18*, 337 – 354.
- Shapiro, A. M. (1999). The relationship between prior knowledge and interactive overviews during hypermedia-aided learning. *Journal of Educational Computing Research, 20*(2), 143-167.
- Shapiro, A. M. (2000). The effect of interactive overviews on the development of conceptual structure in novices learning from hypermedia. *Journal of Educational Multimedia and Hypermedia, 9*(1), 57-78.
- Shepardson, D.P., & Pizzini, E.L. (1991). Questioning levels of junior high-school textbooks and their implications for learning textual information. *Science Education, 75*(6), 673-682.
- Simon, H. A. (1973). The structure of ill structured problems. *Artificial Intelligence, 4*, 181-201.
- Simon, H. A. (1978). Information-processing theory of human problem solving. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes, Vol. 5. Human information processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Slavin, R. J. (1980). Effects of student teams and peer tutoring on academic achievement and time-on-task. *Journal of Experimental Education, 48*, 252-257.
- Slavin, R. J. (1983). When does cooperative learning increase student achievement? *Psychological Bulletin, 94*, 429-445.
- Slavin, R. J. (1986). *Using student team learning* (3<sup>rd</sup> ed.). Baltimore, MD: Johns Hopkins University.

- Slavin, R. J. (1995). *Cooperative learning* (2<sup>nd</sup> ed.). Boston: Allyn & Bacon.
- Slavin, R. J. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*, 43-69.
- Steiner, I. D. (1972). *Group process and productivity*. New York: Academic Press.
- Stevens, R. J., Madden, N. A., Slavin, R. E., & Farnish, A. M. (1987). Cooperative integrated reading and composition: Two field experiments. *Reading Research Quarterly, 22*(4), 433-454.
- Suthers, D. D., & Hyndhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *The Journal of the Learning Sciences, 12*(2), 183-218.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295-312.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics (3rd ed.)*. New York: Harper Collins.
- Taboada, A., & Guthrie, J. T. (2006). Contributions of student questioning and prior knowledge to construction of knowledge from reading information text. *Journal of Literacy Research, 38*(1), 1-35.
- Tao, P. K. (1999). Peer collaboration in solving qualitative physics problems: The role of collaborative talk. *Research in Science Education, 29*(3), 365-383.
- Tao, P. K., & Gunstone, R. F. (1999). Conceptual change in science through collaborative learning at the computer. *International Journal of Science Education, 21*(1), 39-57.

- Teasley, S. D. (1995). The role of talk in children's peer collaborations. *Developmental Psychology, 31*(2), 207-220.
- Thorndike, R. L. (1938). On what type of task will a group do well? *Journal of Abnormal and Social Psychology, 33*, 409-413.
- Towle, A. (2000). *Modern biology*. Austin, TX: Holt, Reinhart, & Winston.
- Tudge, J. (1989). When collaboration leads to regression: Some negative consequences of socio-cognitive conflict. *European Journal of Social Psychology, 19*(2), 123-138.
- Tudge, J. (1992). Processes and consequences of peer collaboration: A Vygotskian analysis. *Child Development, 63*(6), 1364-1379.
- Tudge, J., & Rogoff, B. (1999). Peer influences on cognitive development: Piagetian and Vygotskian perspective. In P. Lloyd and C. Fernyhough, *Lev Vygotsky: Critical assessments: The zone of proximal development, Vol. 3* (pp. 32-56). Florence, KY: Taylor & Francis/ Routledge.
- Van Boxtel, C., & Roelofs, E. (2001). Investigating the quality of student discourse: What constitutes a productive student discourse? *Journal of Classroom Interaction, 37*(1), 55-62.
- Van der Broek, P., Tzeng, Y., Ridsen, K., Trabasso, T., and Basche, P. (2001). Inferential questioning: Effects on comprehension of narrative texts as a function of grade and timing. *Journal of Educational Psychology, 93*(3), 521-529.
- Van Gog, T. V., Ericsson, K. A., Rikers, R., & Paas, F. (2005). Instructional design for advanced learners: Establishing connections between the theoretical frameworks of cognitive load and deliberate practice. *Educational Technology and Development, 53*(3), 73-81.

- Vygotsky, L.S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.
- Wang, T., & Andre, T. (1991). Conceptual change text versus traditional text and application questions versus no application questions in learning about electricity. *Contemporary Educational Psychology, 16*, 103-116.
- Webb, N. M. (1983). Predicting learning from student interaction: Defining the interaction variable. *Educational Psychologist, 18*, 33-41.
- Webb, N. M. (1984). Sex differences in interaction and achievement in cooperative small groups. *Journal of Educational Psychology, 76*, 33-43.
- Webb, N. M. (1989). Peer interaction and learning in small groups. *International Review of Educational Research, 13*, 21-40.
- Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal of Research in Science Teaching, 22*, 366-369.
- Webb, N. M. (1992). Testing a theoretical model of student interaction and learning in small groups. In R. Hertz-Lazarowitz, & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp.102-119). New York: Cambridge University Press.
- Webb, N. M., Nemar, K., M., Chizhik, A. W., & Sugrue, B. (1998). Equity issues in collaborative group assessment: Group composition and performance. *American Educational Research Journal, 35*(4), 607-651.
- Webb, N. M., & Palinscar, A. S. (1996). Group processes in the classroom. In D. C. Berliner and R. C. Calfee (Eds). *Handbook of educational psychology*, (pp. 841-873). New York: Macmillan.

- Wiley, J., & Bailey, J. (2006). Effects of collaboration and argumentation on learning from web pages. In A. M. O'Donnell, C. E. Hmelo-Silver, & G. Erkens, *Collaborative Learning, Reasoning, and Technology* (pp. 297-321). Mahwah, NJ: Lawrence Erlbaum Associates.
- Williams, M. (1996). Student control and instructional technologies. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 957-983). New York: Macmillan.
- Winne, P. H. (1979). Experiments relating teachers' use of higher cognitive questions to student achievement. *Review of Educational Research, 39*, 13-50.
- Winne P. H. (2001). Self-regulated learning viewed from models of information processing. In B. Zimmerman, & D. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp.153-189). Mahwah, NJ: Erlbaum.
- Winters, F. I., & Azevedo, R. (2005). High-school students' regulation of learning during computer-based science inquiry. *Journal of Educational Computing Research, 33*(2), 189-217.
- Winters, F. I., Greene, J. A., & Costich, C. M. (2008). Self-regulation of learning within computer-based learning environments: A critical analysis. *Educational Psychology Review, 20*(4), 429-444.
- Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist, 24*(4), 345-376.
- Wong, B. Y. (1985). Self-questioning instructional research: A review. *Review of Educational Research, 55*(2), 227-268.

- Wood, E., Pressley, M. & Winne, P. H. (1990). Elaborative interrogation effects on children's learning of factual content. *Journal of Educational Psychology*, 82, 741-748.
- Yeh, S. W., & Lehman, J. D. (2001). Effects of learner control and learning strategies on English as a foreign language (EFL) learning from interactive hypermedia lessons. *Journal of Educational Multimedia and Hypermedia*, 10(2), 141-159.
- Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning: From teaching to self-reflective practice* (pp.1-19). New York: Guilford Press.
- Zimmerman, B. J. (2000). Attaining self-regulation: A social-cognitive perspective. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-41). San Diego: Academic Press.