Self-regulated learning with hypermedia: The role of prior domain knowledge

Daniel C. Moos a,*, Roger Azevedo b

a University of Maryland, Department of Human Development, 3304 Benjamin Building, College Park, MD 20742, USA
b Department of Psychology, Institute for Intelligent Systems, University of Memphis 3693 Norriswood Ave., Memphis, TN 38152-6400, USA

Available online 27 April 2007

Abstract

Think-aloud and pre-test data were collected from 49 undergraduates with varying levels of prior domain knowledge to examine the relationship between prior domain knowledge and self-regulated learning with hypermedia. During the experimental session, each participant individually completed a pretest on the circulatory system, and then one 40-min hypermedia learning task during which he or she learned about the circulatory system. Think-aloud data were collected during the 40-min learning task to measure each participant’s use of specific self-regulated learning processes related to planning, monitoring, and strategy use. Results indicate that prior domain knowledge is significantly related to how the participants self-regulated their learning during the 40-min learning task with hypermedia. Specifically, prior domain knowledge is positively related to participants’ monitoring and planning and negatively related to their use of strategies during the hypermedia learning task.

Keywords: Prior domain knowledge; Self-regulated learning; Hypermedia; Cognitive processes; Science; Mental models; Mixed methodology

This study was partially supported by a doctoral fellowship from the University of Maryland awarded to the first author and by funding from the National Science Foundation (Early Career Grant REC#0133346 and REC#0633918) awarded to the second author. The authors wish to acknowledge and thank Jeffrey Greene, Fielding Winters, and Jennifer Cromley for their assistance with the data collection and analysis, and feedback on a previous version of this manuscript. Additionally, the authors wish to acknowledge the feedback given by the anonymous reviewers.

* Corresponding author.
E-mail address: danmoos@msn.com (D.C. Moos).

0361-476X/$ - see front matter © 2007 Elsevier Inc. All rights reserved. doi:10.1016/j.cedpsych.2007.03.001
1. Introduction

In order to explain the variability in domain learning, research in the field of educational psychology has considered a number of factors related to this complex process. This line of research has given rise to several theoretical frameworks, many of which address the relationship between prior domain knowledge and learning (Alexander, 2003; Alexander & Jetton, 2003; Alexander, Jetton, & Kulikowich, 1995; Alexander, Kulikowich, & Schulze, 1994; Alexander & Murphy, 1998; Dochy & Alexander, 1995). Within the field of reading comprehension, for example, research using Kintsch’s Construction Integration model (CI model; 1998) has demonstrated that students with domain-specific prior knowledge are better equipped to understand challenging topics, especially science related topics presented in textbooks (McNamara, Kintsch, Songer, & Kintsch, 1996). According to the CI model, successful understanding of text occurs when the reader develops a text-level representation through integration of individual sentences (Kintsch, 1998). However, science textbooks often include text that has conceptual gaps (Best, Rowe, Ozuru, & McNamara, 2005), and thus students are required to generate inferences in order to develop text-level representations (McNamara, 2004). Students with higher prior domain knowledge are better equipped to make such inferences because they can readily access relevant knowledge structures (Best et al., 2005). Other research has used the Model of Domain Learning (MDL) to examine the differences between learners with limited subject-matter knowledge and more competent learners (Alexander et al., 1995; Murphy & Alexander, 2002). Research using this theoretical approach to domain learning has found that learners with more developed subject-matter knowledge tend to use advanced strategies during learning, while learners with limited subject-matter knowledge tend to approach learning by initially developing a conceptual framework of the domain (Alexander et al., 1995; Murphy & Alexander, 2002).

While the relationship between prior domain knowledge and learning processes has received considerable empirical and theoretical attention (e.g., Afflerbach, 1986; Alexander & Jetton, 2000; Alexander & Kulikowich, 1991; Chi, Feltovich, & Glaser, 1981; Lundeberg, 1987), examining this relationship in the context of computer-based learning environments (CBLEs) is an emerging area of study (Azevedo, 2005; Graesser, McNamara, & VanLehn, 2005; Shapiro, 1999, 2004; Shapiro & Niederhauser, 2004). Research in this area is important because though CBLEs are seen as promising educational tools, the theoretical understanding of how students learn with these environments is still developing (Winne, 2005). In order to advance the theoretical understanding of the process involved in learning with CBLEs, research is needed that empirically examines factors that have been traditionally related to academic learning, such as prior domain knowledge. Several distinct CBLEs exist, and thus it is important to systematically examine the role of prior domain knowledge in a variety of CBLEs that have been used in the classroom. In particular, hypermedia is one type of CBLE that has been used extensively to help students learn about challenging topics (Azevedo & Cromley, 2004; Azevedo, Cromley, & Seibert, 2004; Azevedo, Cromley, Winters, Moos, & Greene, 2005; Azevedo, Guthrie, & Seibert, 2004; Jacobson & Archolidou, 2000; Lajoie & Azevedo, 2006; Shapiro, 1999, 2000, 2004).

Hypermedia has several defining characteristics which make it distinct from other CBLEs. Nodes, which can consist of text, audio, and/or animation, are the basic unit of hypermedia environments (Jonassen & Reeves, 1996). When learning with hypermedia, students can access any node and in multiple sequences, depending on various factors such
learning goals and prior domain knowledge. Because the presentation of non-linear information offers a certain degree of navigational control over the sequencing of instructional content, students can actively participate in the learning process (Jonassen & Reeves, 1996; Williams, 1996). However, empirical research has suggested that the effectiveness of hypermedia environments may be mediated by several factors, including prior domain knowledge (see Chen, Fan, & Macredie, 2006; Shapiro & Niederhauser, 2004 for recent reviews). Research has demonstrated that students with higher prior domain knowledge learn differently with hypermedia environments than students with lower prior domain knowledge.

2. Prior domain knowledge and learning with hypermedia

Research indicates that prior domain knowledge is related to how students learn with hypermedia, particularly with how they navigate within this learning environment (Caliser & Gurel, 2003; McDonald & Stevenson, 1998; Shin, Schallert, & Savenye, 1994). Students need to manage a high degree of control in order to effectively navigate through hypermedia (Lawless & Brown, 1997) because this environment allows students to choose which information to access (Azevedo, 2005; Shapiro, 1999; Williams, 1996). As suggested by previous research, some students may have difficulty navigating in hypermedia due to limited prior domain knowledge (Chen & Ford, 1998; Last, O’Donnell, & Kelly, 2001; Nelson, 2000). Chen et al. (2006) suggest that navigation within hypermedia is, in part, dependent on an understanding of the conceptual structure of the domain. This understanding guides students’ interaction with the non-linear nature of hypermedia. Thus, students who have limited understanding of the conceptual structure of the domain have little to guide their interaction with hypermedia, which explains why students with lower prior domain knowledge have more difficulty navigating in this environment (Shapiro, 2004).

This line of research has advanced the field by examining the relationship between prior domain knowledge and navigation during learning with hypermedia. However, while navigation is clearly a critical component of learning with hypermedia, navigational choices do not provide traces (i.e., observable indicators) of other factors that have been shown to be related to learning with hypermedia, such as self-regulated learning (Azevedo, 2005; Winne, 2005; Winne & Hadwin, 1998; Winne & Perry, 2000). Recent research has found that when students are learning in an environment with multiple representations and non-linear information they need to self-regulate certain aspects of their learning (Azevedo & Cromley, 2004). As such, this line of research has used the self-regulated learning theory to examine how students learn with hypermedia (e.g., Moos & Azevedo, 2006).

3. Self-regulated learning theory (SRL)

Self-regulated learning (SRL) involves actively constructing an understanding of a topic/domain by using strategies and goals, regulating and monitoring certain aspects of cognition, behavior, and motivation, and modifying behavior to achieve a desired goal (Pintrich, 2000). Though this definition of SRL is commonly used, the field of SRL consists of various theoretical perspectives that sometimes focus on different constructs (Boekaerts, Pintrich, & Zeidner, 2000; Zimmerman & Schunk, 2001). This study draws from Winne (2001) and Winne and Hadwin’s (1998) information processing theory
This IPT theory suggests a 4-stage model of self-regulated learning. In the first phase, the learner constructs a perception of the task. These perceptions are drawn from several sources, including information provided in the outside environment (i.e. task conditions) and information the learner retrieves from long-term memory (i.e. cognitive conditions). According to the IPT theory, prior domain knowledge is a fundamental source of information when a learner constructs a perception of the task. Prior domain knowledge, retrieved from long-term memory (Ericsson & Kintsch, 1995; Sweller, 2003, 2004), facilitates the learner’s definition of the task and task performance (Winne, 2001).

In phase two, the learner frames multifaceted goals and plans how to reach the goal(s) (Butler & Winne, 1995; Winne & Hadwin, 1998). In phase three, the learner enacts tactics and/or strategies to meet these goals. Lastly, in phase four, the learner may make adaptations to schemas that structure various self-regulated processes. However, an underlying assumption of this theory is that SRL is recursive because of a feedback loop. That is, information processed in one phase can become an input to subsequent information processing. Additionally, metacognitive monitoring is the key to self-regulated learning according to this theoretical approach to SRL (Butler & Winne, 1995; Winne, 1996, 1997). Monitoring is defined as processes that compare two chunks of information (i.e. learning goal and current domain knowledge; Winne, 2001). Metacognitive monitoring produces information that allows learners to determine if there is a discrepancy between any goals and their current level of domain knowledge. If monitoring does reveal a discrepancy, then learners may adapt their planning and/or strategies in order to meet the goal. This assumption suggests that adaptation is dependent on metacognitive monitoring. Empirical research has identified a number of specific metacognitive processes that are related to learning, such as Feeling of Knowing (FOK) and Judgment of Learning (JOL). FOKs are defined as an awareness of having read something in the past and having some understanding of it but not begin able to recall it on demand, while JOLs are defined as an awareness of not knowing or understanding everything read (Azevedo et al., 2005). While the IPT theory of SRL highlights the role of strategy use, monitoring, and planning in learning, this theory also indicates that prior domain knowledge is a critical component to domain learning.

4. Prior domain knowledge and self-regulated learning with hypermedia

According to the IPT theory, working memory (i.e. hypothetical location where information becomes a topic of information processing) has a limited capacity of three chunks when tasks are complicated (Sweller, van Merrienboer, & Pass, 1998). When students encounter a task in which they have limited prior domain knowledge, the majority of this capacity may be used for processing information. Thus, in this case, students may not be able to use a large variety of SRL processes because there is limited remaining capacity in working memory. This assumption provides a theoretical explanation of why students with lower prior domain knowledge may use few monitoring and planning processes during learning, and instead rely on a small subset of strategies. However, little empirical research has used the IPT theory of SRL to empirically explain how students with varying levels of prior domain knowledge self-regulate their learning with hypermedia (Winne, 2001). In order to extend the IPT approach to SRL, empirical research is needed which examines the relationship between prior domain knowledge and use of key SRL processes, including strategies, monitoring, and planning processes. In order to address this issue,
this study attempts to empirically identify this relationship by using a think-aloud protocol to examine college students’ prior domain knowledge and their use of strategies, monitoring, and planning processes during learning with hypermedia.

5. Current study

College students were of research interest for this study because a sample was needed that had the potential to have variability in both prior domain knowledge and use of specific SRL processes during learning with hypermedia. In other words, if the majority of participants in the sample either had low prior domain knowledge and/or used very few SRL processes during learning with hypermedia, then the variability would be minimal and the power to detect the relationship between prior domain knowledge and use of SRL processes would be limited. Previous research has indicated that grade-level middle school students may infrequently use SRL processes, and that even gifted middle school students may not use many SRL processes related to strategies and metacognitive processes (Carr, Alexander, & Schwanenflugel, 1996; Greene, Moos, Azevedo, & Winters, in press; Zimmerman & Martinez-Pons, 1986). Furthermore, research has also demonstrated that many high school students have limited knowledge of challenging science related topics (the domain in this study) and that they have difficulty using CBLEs to learn about these topics (Azevedo, Winters, & Moos, 2004). On the other hand, previous research has demonstrated that while some college students have limited prior domain knowledge of challenging science topics and use few SRL processes during learning (Moos & Azevedo, 2006), other college students demonstrate high prior domain knowledge and use a variety of SRL processes during learning with hypermedia (Azevedo & Cromley, 2004). As such, the developmental group of college students was the most appropriate sample for this study.

The objective of this study was to empirically examine the relationship between prior domain knowledge and college students’ use of specific SRL processes (strategy use, monitoring, and planning) while learning with hypermedia. In this study, prior domain knowledge was defined as conceptual understanding of the circulatory system, which integrates declarative, procedural, and inferential knowledge (Chi, 2000, 2005). Converging process data from think-aloud methodology (Ericsson, 2006; Ericsson & Simon, 1993) and product data from a pretest, this study extends previous research examining factors related to learning with hypermedia. Specifically, this study addressed was the following two research questions:

(1) To what extent is students’ prior domain knowledge related to how they self-regulate their learning with hypermedia?

(2) How do verbalizations, as captured by think-alouds, demonstrate the relationship between students’ prior domain knowledge and their self-regulation of learning with hypermedia?

Based on previous research, the following three hypotheses were proposed:

1. Research using the MDL to examine the differences between students with limited subject-matter knowledge and more competent students has found that prior domain knowledge is related to the use of strategies (Alexander et al., 1995; Murphy &
Alexander, 2002). Specifically, students with higher prior domain knowledge tend to use a wider variety of more advanced strategies, while students with lower prior domain knowledge tend to use less varied strategies. Based on this line of research, it was expected that students’ prior domain knowledge would be positively correlated with their use of strategies during learning with hypermedia.

2. The IPT theory assumes that a deep understanding of the conceptual structure of the domain facilitates self-regulated learning, including goal-setting and developing plans to meet these goals (Winne, 2001). Thus, it was hypothesized that prior domain knowledge would be positively correlated with the use of planning processes during learning with hypermedia.

3. Previous research also indicates that students’ navigation is related to their understanding of the conceptual structure of the domain (Shapiro, 2004). In order to effectively navigate in hypermedia, students need to be able to monitor both their emerging understanding and the relevancy of the content. However, limited conceptual knowledge of the domain may limit students’ ability to monitor the relevancy of the content. As such, it was expected that students’ prior domain knowledge would be positively correlated with their monitoring activity during learning with hypermedia.

6. Method

6.1. Participants

Forty-nine (n = 49) undergraduate students enrolled in educational psychology classes from a large Mid-Atlantic public university participated in this study. The participants received extra credit in their classes for participating in this study. Their average age was 21.39 (SD = 6.17); there were 42 women (86%) and 7 men (14%). This sample consisted primarily of females, which is typical of most education classes. The sample included 12 freshmen (24%), 5 sophomores (10%), 19 juniors (39%), and 13 seniors (27%). All participants were individually tested in a research laboratory associated with the Human Development program at the Mid-Atlantic public university.

6.2. Measures

Each participant individually completed the following paper-and-pencil materials: A consent form, a participant questionnaire, and a pretest. All of these measures have been used in previously published studies examining how students use hypermedia to learn about the circulatory system (e.g., Azevedo et al., 2005). The participant questionnaire collected information about the participants’ gender, age, and any relevant work experience with biology. The pretest is comprised of three sections: (a) a matching task in which the participants were asked to match 13 words with their corresponding definitions, (b) a labeling task in which the participants were asked to label 10 parts of the heart, and (c) an essay which asked the participants to, “Please write down everything you can about the circulatory system. Be sure to include all the parts and their purpose, explain how they work both individually and together, and also explain how they contribute to the healthy functioning of the body”. The pretest is designed to measure participants’ understanding of the circulatory system.
6.3. Materials

6.3.1. Hypermedia environment

The hypermedia environment used in this study was a commercially-based electronic encyclopedia, Microsoft Encarta Reference Suite™ (2003). Participants used Encarta to learn about the circulatory system and had access to three relevant articles (i.e., circulatory system, blood, and heart), which together were comprised of 16,900 words, 35 illustrations, 107 hyperlinks, and 18 sections and one 2-min animation of diastole and systole. Participants could freely use all of these articles, hyperlinks, and video during the learning task.

6.4. Procedure

Participants were all individually tested in a university laboratory. First, participants were given as much time as needed to complete the consent form and the participant questionnaire. Second, participants were given 20 min to complete the pretest. Third, the researcher provided instructions for the learning task: “You are being presented with a hypermedia encyclopedia, which contains textual information, static diagrams, and a digitized video clip of the circulatory system. We are trying to learn more about how students use hypermedia environments to learn about the circulatory system. Your task is to learn all you can about the circulatory system in 40 min. 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. We ask you to ‘think aloud’ continuously while you use the hypermedia environment to learn about the circulatory system. I’ll be here in case anything goes wrong with the computer or the equipment. Please remember that it is very important to say everything that you are thinking while you are working on this task”.

After receiving the instructions, participants were given a 5-min training session of the hypermedia environment in which the most relevant articles for the circulatory system were identified. They also practiced navigating (e.g., forward and backward buttons), using the search tools (e.g., find function), and accessing multiple representations (text, static diagrams, and digitized video clip). Next, the participants were given 40 min to learn about the circulatory system with the hypermedia environment. The researcher remained nearby to remind participants to keep verbalizing when they were silent for more then three seconds (e.g., “Say what you are thinking”). On average, the researcher needed to remind participants to think-aloud seven times during the 40-min learning task. The participants’ verbalizations during the 40-min learning task were recorded and later used to analyze their self-regulated learning. This pretest protocol, in which the participant is given 20 min to complete the pretest, five minutes for the training session of the hypermedia environment, and 40 min to learn about the circulatory system, is based on extensive research that has examined how undergraduates self-regulate their learning of the circulatory system with hypermedia (e.g., Azevedo et al., 2005; Azevedo, Cromley, et al., 2004; Azevedo, Guthrie, et al., 2004; Moos & Azevedo, 2006). This line of research, which has collected product and process data from over 500 undergraduate students, has demonstrated that 20 min is a sufficient amount of time for participants to complete the pretest measure, and the mean pretest completion time for participants in this study supports this assertion ($M_{time} = 16.07$). Additionally, previous research has demonstrated that undergraduates are able to shift from a low mental model to a high mental model of
the circulatory system after using a hypermedia environment for 40 min (Azevedo et al., 2005; Greene & Azevedo, 2006).

6.5. Coding and scoring

In this section, the coding and scoring of the participants’ pretest (matching, labeling, and mental model essay) are discussed. In addition, the coding scheme used to analyze the participants’ use of self-regulatory processes during learning is provided. Finally, inter-rater agreement is discussed.

6.5.1. Matching task

For the matching task, a participant received either a 1 (for a correct match between a concept and its corresponding definition) or a 0 (for an incorrect match between a concept and definition) for each item on his or her pretest (possible range 0–13).

6.5.2. Labeling task

For the labeling task, a participant received either a 1 (for a correct labeling of one part of the heart) or a 0 (for an incorrect or blank labeling for one part of the heart) for each item on his or her pretest (possible range 0–10).

6.5.3. Mental model essay

The third part of the pretest was the mental model essay. We scored the participants’ essay on the circulatory system using Azevedo and colleagues’ method (Azevedo & Cromley, 2004; Azevedo et al., 2005; Azevedo, Cromley, et al., 2004; Azevedo, Guthrie, et al., 2004 which is based on Chi and colleagues’ research (Chi, 2000, 2005; Chi, de Leeuw, Chiu, & LaVancher, 1994). Azevedo and colleagues’ scheme for scoring the essay ranges from 1 to 12, and represents the progression from no understanding to the most accurate understanding of the circulatory system: (1) no understanding, (2) basic global concept, (3) basic global concept with purpose, (4) basic single loop model, (5) single loop with purpose, (6) advanced single loop model, (7) single loop model with lungs, (8) advanced single loop model with lungs, (9) double loop concept, (10) basic double loop model, (11) detailed double loop model, and (12) advanced double loop model. A complete description of the necessary features for each mental model is provided in Table 1.

6.5.4. Self-regulatory processes

A think-aloud protocol methodology (Ericsson, 2006; Ericsson & Simon, 1993) was used to capture participants’ use of SRL processes during learning. The think aloud protocol is a trace methodology that has been recently used to examine SRL processes during learning with hypermedia environments (see Azevedo, 2005). The think aloud has an extensive history in cognitive psychology and cognitive science (see Ericsson, 2006; Ericsson & Simon, 1993; Newell & Simon, 1972 for extensive reviews). As suggested by Anderson (1987), cognitive psychology and cognitive science have used both concurrent and retrospective think aloud protocols to measure cognitive processes in a multitude of tasks including medical diagnosis, electronics troubleshooting, computer programming, and mathematical reasoning. (Chi, Glaser, & Farr, 1988; Ericsson, 1996, 2006; Ericsson & Smith, 1991; Feltovich, Ford, & Hoffman, 1997; Laszlo, Meutsch, & Viehoff, 1988; Norris, 1991). In this study, concurrent think alouds were used, which assume that thought
Table 1
Necessary features for each type of mental model (from Azevedo and Cromley, 2004)

1. **No understanding**
2. **Basic global concepts**
   - Blood circulates
3. **Global concepts with purpose**
   - Blood circulates
   - Describes “purpose”—oxygen/nutrient transport
4. **Single loop—basic**
   - Blood circulates
   - Heart as pump
   - Vessels (arteries/veins) transport
5. **Single loop with purpose**
   - Blood circulates
   - Heart as pump
   - Vessels (arteries/veins) transport
   - Describe “purpose”—oxygen/nutrient transport
6. **Single loop—advanced**
   - Blood circulates
   - Heart as pump
   - Vessels (arteries/veins) transport
   - Describe “purpose”—oxygen/nutrient transport
   - Mentions one of the following: electrical system, transport functions of blood, details of blood cells
7. **Single loop with lungs**
   - Blood circulates
   - Heart as pump
   - Vessels (arteries/veins) transport
   - Mentions lungs as a “stop” along the way
   - Describe “purpose”—oxygen/nutrient transport
8. **Single loop with lungs—advanced**
   - Blood circulates
   - Heart as pump
   - Vessels (arteries/veins) transport
   - Mentions Lungs as a “stop” along the way
   - Describe “purpose”—oxygen/nutrient transport
   - Mentions one of the following: electrical system, transport functions of blood, details of blood cells
9. **Double loop concept**
   - Blood circulates
   - Heart as pump
   - Vessels (arteries/veins) transport
   - Describes “purpose”—oxygen/nutrient transport
   - Mentions separate pulmonary and systemic systems
   - Mentions importance of lungs
10. **Double loop—basic**
    - Blood circulates
    - Heart as pump
    - Vessels (arteries/veins) transport
    - Describe “purpose”—oxygen/nutrient transport
    - Describes loop: heart-body-heart-lungs-heart
processes are a sequence of states, and that information in a state is relatively stable (Ericsson & Simon, 1993). This assumption asserts that verbalizations uttered by participants during learning do not disrupt the learning process, nor do they alter the cognitive processes underlying task performance. It is important to highlight the fact that “subjects verbalizing their thoughts while performing a task do not describe or explain what they are doing” (Ericsson & Simon, 1993, pg. xiii) in a concurrent think-aloud protocol. If subjects are not asked to reflect, describe, and/or explain their thoughts during learning, but rather are asked to simply verbalize thoughts entering their attention, then it is assumed that the sequence of thoughts will not be disrupted. Empirical evidence has supported this assertion. For example, Deffner (1989), Heydemann (1986), & Rhenius & Heydemann (1984) all found that asking participants to think aloud during learning did not disrupt cognitive processes, as reflected in the participants’ performance in these studies.

While the think aloud protocol has been most popular in reading comprehension (Pressley & Afflerbach, 1995), it has been shown to be an excellent tool in gathering verbal accounts of SRL and mapping out the use of self-regulatory processes during learning (e.g., Azevedo, 2005; Azevedo & Cromley, 2004). Furthermore, research has used the concurrent think aloud protocol to examine learning processes with hypermedia. For example, Azevedo, Guthrie, et al. (2004) used the concurrent think aloud protocol to examine the role of SRL in fostering students’ conceptual understanding of complex systems while using hypermedia. Other studies also support the effectiveness of the concurrent think-aloud in measuring SRL as an event, including Azevedo, Winters, et al. (2004) & Moos & Azevedo (2006). In this line of research, the concurrent think aloud was used to examine how students plan, monitor, use strategies, and handle task difficulties while learning about challenging science-related topic with hypermedia. In sum, the proven capacity of the concurrent think-aloud protocol to capture learning processes in a dynamic learning situation provides support for the use of this protocol to measure SRL (Winne & Perry, 2000).

A coding model developed by Azevedo and colleagues (2005) was used to code the participants’ verbalizations gathered from the concurrent think-aloud protocol. Their model was based on several recent models of SRL (Pintrich, 2000; Winne, 2001; Winne & Hadwin, 1998; Winne & Perry, 2000; Zimmerman, 2000, 2001). This model includes key components of Pintrich’s (2000) formulation of self-regulation as a four-phase process

<table>
<thead>
<tr>
<th>Table 1 (continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11. <strong>Double loop—detailed</strong></td>
</tr>
<tr>
<td>• Blood circulates</td>
</tr>
<tr>
<td>• Heart as pump</td>
</tr>
<tr>
<td>• Vessels (arteries/veins) transport</td>
</tr>
<tr>
<td>• Describe “purpose”—oxygen/nutrient transport</td>
</tr>
<tr>
<td>• Describes loop: heart-body-heart-lungs-heart</td>
</tr>
<tr>
<td>• Structural details described: names vessels, describes flow through valves</td>
</tr>
<tr>
<td>12. <strong>Double loop—advanced</strong></td>
</tr>
<tr>
<td>• Blood circulates</td>
</tr>
<tr>
<td>• Heart as pump</td>
</tr>
<tr>
<td>• Vessels (arteries/veins) transport</td>
</tr>
<tr>
<td>• Describe “purpose”—oxygen/nutrient transport</td>
</tr>
<tr>
<td>• Describes loop: heart-body-heart-lungs-heart</td>
</tr>
<tr>
<td>• Structural details described: names vessels, describes flow through valves</td>
</tr>
<tr>
<td>• Mentions one of the following: electrical system, transport functions of blood, details of blood cell</td>
</tr>
</tbody>
</table>
and extends these components to capture the major phases of self-regulation. The coding scheme used in this study was modified from Azevedo and colleagues (2005) to capture the SRL processes in this particular study, and includes 27 SRL processes from the three SRL categories of planning, monitoring, and strategy use (see Appendix A for the list of SRL codes). The planning category consists of planning, prior domain knowledge activation, recycling goals into working memory, and setting sub-goals. The monitoring category consists of content evaluation, feeling of knowing (FOK), identifying adequacy of information, judgment of learning (JOL), monitoring progress towards goal, monitoring use of strategies, and self-questioning. The strategy use category consists of coordinating informational sources, drawing, finding locating in environment, free searching, goal-directed searching, hypothesizing, making inferences, knowledge elaboration, memorizing, using mnemonics, reading new paragraph, reading notes, re-reading, selecting new informational source, summarizing, and taking notes.

The raw data collected from this study consisted of 1960 min (32.7 h) of audio recordings from 49 participants who gave extensive verbalizations while learning about the circulatory system. During the first phase of this analysis, the audio tapes were transcribed and a text file was created for each student. This phase of the analysis yielded a total of 575 double-spaced pages ($M = 11.73$ pages per participant). All of the transcriptions were then coded by assigning one of the SRL variables to each segment (see Appendix B for an example of a coded page from a transcript). It should be noted that the SRL coding scheme was not designed to segment and code all of the participants’ verbalizations. If a verbalization was not codeable (i.e. the content of the verbalization was not related to any of the SRL codes in the coding scheme), then the segment was not assigned a SRL code. For example, the segments in which UG114 was verbalizing her reading (as presented in Appendix B) were not coded because reading is not related to any of the SRL codes in the coding scheme.

The coding phase yielded a total of 2697 coded SRL segments for all participants ($M_{SRL} = 55.04$ per participant). After coding each transcript, the individual SRL codes were then collapsed into one of the three corresponding SRL categories (planning, monitoring, or strategy use). These raw SRL frequencies were then converted into proportions in order to control for the individual differences in the extent to which each participant verbalized during the 40-min hypermedia learning task. The proportion of SRL processes for each category and for each participant was calculated by dividing the frequency of the individual SRL processes by the total frequency of all SRL processes for each individual participant. For example, participant UG052 had a total of 44 coded SRL processes. Thirty of these 44 codes were coded as strategy use. Thus, the proportion of strategy use codes for participant UG052 is 30/44, or 68%. In terms of the remaining SRL codes for participant UG052, 16% ($n = 7$) were monitoring and 16% ($n = 7$) were planning.

6.6. Inter-rater agreement

Inter-rater reliability was established for coding of the participants’ SRL by comparing the individual coding of a graduate student, who was trained to use an adapted version of Azevedo and colleagues (2005) coding scheme, with that of the second author (for complete details of coding scheme, see Azevedo et al., 2005). Fifty-one percent of the transcripts ($n = 25$) were used for inter-rater reliability, and there was agreement on 1642 out of 1685 coded SRL segments, yielding a reliability coefficient of .97. In addition, inter-rater reliability was also established for the scoring of the pretest. For the essays, there was agreement on 93
out of a total of 98 scored essays, yielding an inter-rater agreement of .95. Disagreements on
the SRL coding and scoring of the pretest essays were resolved through discussion.

7. Results

In order to address the first research question, a principal component analysis (PCA)
was first run on the three sections of the pretest to determine if these sections truly measure
one factor, general domain knowledge of the circulatory system. PCA is a common statistical
technique that allows for parsimonious reduction of the data (Guadagnoli & Velicer,
1988; Velicer & Jackson, 1990). In terms of the minimum sample size for PCA, some
researchers have suggested that the participants to item ratio should be 10:1 (Nunnally,
1978), and recent research has supported using this ratio to determine the minimum sam-
ple size (Osborne & Costello, 2004). Because three items were used in this PCA (labeling,
matching, and essay scores), a minimum sample of 30 participants was required. The final
sample size of this study (N = 49) exceeded this required minimum. Additionally, the Bart-
lett Test of Sphericity (Bartlett, 1951) χ² (3, N = 49) = 27.254, p < .001 and the Kaiser–
Meyer–Olkin (KMO) value of .62 indicated that there was adequate sampling and suitable
 correlation matrix for this factor analysis (see Table 2 for correlation matrix of labeling,
matching, and essay). The KMO value of .62 exceeded the minimum value of .50 for a satis-
factory PCA to proceed (Tabachnick & Fidell, 1996).

The Kaiser criterion rule (K1), in which factors with eigenvalues greater than one are
retained, and Cattell’s subjective scree test (1966) were used to determine the number of fac-
tors to retain. Both the K1 extraction criteria and the Catell’s subjective scree test suggested
that one factor should be retained. Additionally, this one factor accounted for 62.47 percent
of the variance. Lastly, all three sections (matching, labeling, and essay) loaded highly on one
factor (loadings ≥ .60; see Table 3). These criteria indicated that the three sections of the pre-
test measure one factor. Thus, we labeled this one factor as conceptual knowledge of the cir-
culatory system, which in this study includes two measures of declarative knowledge and one
measure of mental model of the circulatory system (Chi, 2000, 2005; Chi et al., 1994).

The first research question addressed the relationship between prior domain knowledge
and how students regulate their learning with hypermedia. To determine this relationship,
correlations between prior domain knowledge (as indicated by participants’ factor score
from the PCA) and proportion of SRL processes used during learning (planning, monitor-
ing, and strategy use) were calculated. The correlation analyses revealed significant rela-
tionships between prior domain knowledge and use of specific SRL processes. As
shown in the correlation matrix in Table 4, there were significant correlations between
prior domain knowledge and planning (r = .37, p < .01), monitoring (r = .30, p < .05),
and strategy use (r = −.37, p < .01). These correlations analyses indicated that participants
with higher prior domain knowledge of the circulatory system tended to use significantly

<table>
<thead>
<tr>
<th>Correlations between sections on pretest</th>
<th>Labeling</th>
<th>Essay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matching task</td>
<td>.55**</td>
<td>.30*</td>
</tr>
<tr>
<td>Labeling task</td>
<td>—</td>
<td>.45**</td>
</tr>
</tbody>
</table>

* p < .05.

** p < .01 (two-tailed).
more planning and monitoring processes during learning, while participants with lower prior domain knowledge tended to use significantly more strategies during learning with hypermedia.

Examining the raw frequencies of individual SRL processes may further explain the relationship between prior domain knowledge and self-regulated learning (see Table 5 for total raw frequencies of individual SRL processes, by prior domain knowledge). These raw frequencies suggest that students with higher prior domain knowledge planned their learning by activating prior domain knowledge and recycling goals in working memory more frequently than student with lower prior domain knowledge. Furthermore, these raw frequencies also imply that students with higher prior domain knowledge monitored their learning by evaluating the content in hypermedia and expressing a feeling of knowing more frequently than students with lower prior domain knowledge. On the other hand, students with lower prior domain knowledge used the strategies of note-taking, summarizing, and memorizing more frequently than students with higher prior domain knowledge.

Although the results from the correlation analyses are useful in providing an overview of the relationship between prior domain knowledge and SRL processes, these correlations tell us little about how students with differing levels of prior domain knowledge used these processes during learning. While students with higher prior domain knowledge tended to use monitoring and planning processes during learning, how did they use these processes? Conversely, how did students with lower prior domain knowledge use strategies? Research question two addresses how students used SRL processes during learning by examining particular episodes gathered from the think-aloud protocol (Wells & Arauz, 2006). The following section presents portions of coded transcriptions from four different participants. These coded transcriptions highlight how students with differing levels of prior domain knowledge used SRL processes during learning. The first two coded transcriptions demonstrate how students with higher prior domain knowledge tended to rely on monitoring processes and very few strategies, while students with lower prior domain knowledge tended to primarily use a specific subset of strategies during learning. The third and fourth examples demonstrate the difference in how students use planning processes during learning.
The first transcription is from a student who had high prior domain knowledge of the circulatory system (as indicated by her pretest), and primarily self-regulated her learning by monitoring both her understanding of the circulatory system and relevancy of the environment’s content. The second transcription comes from a student who had low prior domain knowledge (as indicated by his pretest) and primarily self-regulated his learning

<table>
<thead>
<tr>
<th>SRL processes</th>
<th>Low prior domain knowledge (n = 25)</th>
<th>High prior domain knowledge (n = 24)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Planning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Prior domain knowledge activation</td>
<td>29</td>
<td>97</td>
</tr>
<tr>
<td>Recycle goal in working memory</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Sub-goals</td>
<td>54</td>
<td>58</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>88</td>
<td>184</td>
</tr>
<tr>
<td><strong>Monitoring</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content evaluation</td>
<td>54</td>
<td>81</td>
</tr>
<tr>
<td>Feeling of knowing (FOK)</td>
<td>144</td>
<td>158</td>
</tr>
<tr>
<td>Identify adequacy of information</td>
<td>18</td>
<td>23</td>
</tr>
<tr>
<td>Judgment of learning (JOL)</td>
<td>98</td>
<td>88</td>
</tr>
<tr>
<td>Monitoring progress toward goals</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Monitor use of strategies</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>Self-questioning</td>
<td>28</td>
<td>12</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>362</td>
<td>378</td>
</tr>
<tr>
<td><strong>Strategy use</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coordinating informational sources</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>Draw</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Find location in environment</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Free search</td>
<td>12</td>
<td>28</td>
</tr>
<tr>
<td>Goal-directed search</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>Hypothesizing</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Inferences</td>
<td>25</td>
<td>37</td>
</tr>
<tr>
<td>Knowledge elaboration</td>
<td>23</td>
<td>21</td>
</tr>
<tr>
<td>Memorization</td>
<td>22</td>
<td>6</td>
</tr>
<tr>
<td>Mnemonics</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>Read new paragraph</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Read notes</td>
<td>18</td>
<td>30</td>
</tr>
<tr>
<td>Re-reading</td>
<td>116</td>
<td>104</td>
</tr>
<tr>
<td>Selecting new informational source</td>
<td>53</td>
<td>60</td>
</tr>
<tr>
<td>Summarization</td>
<td>301</td>
<td>228</td>
</tr>
<tr>
<td>Taking notes</td>
<td>266</td>
<td>185</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>900</td>
<td>785</td>
</tr>
</tbody>
</table>

*In order to categorize each student as either having high or low prior domain knowledge, a median split was done on the factor score. This median split was not used in a statistical analysis, but rather used to help identify differences in SRL by prior domain knowledge.*
by using strategies. Each excerpt includes the segment number, the student’s think aloud protocol (regular typeface), what the student read from the hypermedia (italics), and the SRL code associated with the segment (refer to Appendix A for a description of SRL codes used by each student). Each coded transcriptions is followed by a summary that further describes how the student self-regulated his or her learning in the particular excerpt.

In both of these coded transcriptions, each participant decided to read the same two passages from the hypermedia environment. One passage introduces the circulatory system, while the other describes the function of the heart. This first transcription is from a participant whose pretest indicated that she had high prior domain knowledge of the circulatory system.

<table>
<thead>
<tr>
<th>Segment</th>
<th>[SRL Process]SRL Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I am going to start with the circulatory system just because I am already there...</td>
</tr>
<tr>
<td>2</td>
<td>...and I’m just reading the introduction...circulatory system...it also known as the cardiovascular system and it deals with the heart...it transports oxygen and nutrients and it takes away waste...</td>
</tr>
<tr>
<td>3</td>
<td>...um, it does stuff with blood and I’m kind of remembering some of this from bio in high school, but not a lot of it, um...</td>
</tr>
<tr>
<td>4</td>
<td>Reads: The heart and the blood and the blood vessels are the three structural elements and the heart is the engine of the circulatory system, it is divided into four chambers.</td>
</tr>
<tr>
<td>5</td>
<td>I knew this one, two right and two left...the atrium, the ventricle and the left atrium, and the left ventricle...</td>
</tr>
<tr>
<td>6</td>
<td>...okay start the introduction [of the heart], just kind of scout it out real quick...and there’s a section called function of the heart...and it looks like it will give me what I need to know...</td>
</tr>
<tr>
<td>7</td>
<td>...um...introduction, oh that’s just basic stuff that we’ve been doing</td>
</tr>
<tr>
<td>8</td>
<td>Reads: Structure of the heart has four chambers</td>
</tr>
<tr>
<td>9</td>
<td>We did that...</td>
</tr>
<tr>
<td>10</td>
<td>Reads: The atria are also known as auricles. They collect blood that pours in from veins.</td>
</tr>
<tr>
<td>11</td>
<td>So, it looks like the first step is atria in the system and then the veins</td>
</tr>
</tbody>
</table>

In segment 1, the student decided to start in the introduction of the circulatory system, and then summarized some of the information provided in this section (segment 2). After a brief summarization, the student monitored her understanding of this information (segment 3). She then followed this monitoring activity with more reading (segment 4). She continued to monitor her understanding of this material by stating that she learned this information in a high school biology class (segment 5). After identifying that she had previously learned this information, she hyperlinked from the circulatory system...
article to the heart article and monitored the relevancy of the content in the environment (segment 6). After identifying that this article may have relevant information, she decided to focus on the function of the heart within this section. In this section, she monitored her understanding (segments 7 and 9) while reading (segments 8 and 10). While she self-regulated her learning by primarily using monitoring processes, she occasionally used strategies during this learning episode (segments 2 and 11).

This second transcription is from a participant whose pretest indicated that he had low prior domain knowledge of the circulatory system.

<table>
<thead>
<tr>
<th>Segment</th>
<th>SRL Process</th>
<th>SRL Category</th>
<th>Strategy</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I am going to the introduction...</td>
<td>No Code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Reads: Circulatory system, or cardiovascular system, in humans, the combined function of the heart, blood, and blood vessels to transport oxygen and nutrients to organs and tissues throughout the body and carry away waste products...</td>
<td>No Code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>I'm going to take notes...transport oxygen...nutrients...to organs...and tissues...and carry away waste products...</td>
<td>Taking Notes</td>
<td>Strategy Use</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Reads: Among its vital functions, the circulatory system increases the flow of blood to meet increased energy demands during exercise and regulates body temperature. In addition, when foreign substances or organisms invade the body, the circulatory system swiftly conveys disease-fighting elements of the immune system, such as white blood cells and antibodies to regions under attack...</td>
<td>No Code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>I'm writing down the structural elements...</td>
<td>Taking Notes</td>
<td>Strategy Use</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Reads: The heart is the engine of the circulatory system. It is divided into four chambers: The right atrium, the right ventricle, the left atrium, and the left ventricle. The walls of the chambers are made of a special muscle called myocardium, which contract continuously and rhythmically to pump blood...</td>
<td>No Code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>...okay, the heart...engine...the chambers...right and left atrium...right and left ventricle, okay...special muscle...myocardium...mmmm...</td>
<td>Taking Notes</td>
<td>Strategy Use</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Reads: The human heart has four chambers, the upper two chambers...the right side of the heart is responsible for pumping oxygen-poor blood to the lungs...This oxygen-poor blood feeds into two large veins, the superior vena cava and inferior vena cava. The right atrium conducts blood to the right ventricle, and the right ventricle pumps blood into the pulmonary artery. The pulmonary artery carries the blood to the lungs, where it picks up a fresh supply of oxygen and eliminates carbon dioxide.</td>
<td>No Code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Reads: This oxygen-poor blood feeds into two large veins, the superior</td>
<td>Re-read</td>
<td>Strategy Use</td>
<td></td>
</tr>
</tbody>
</table>
Okay, superior and inferior which empty it into right atrium of the heart... the right atrium conducts blood to the right ventricle... and the right ventricle pumps blood into the pulmonary artery...

**Reads:** The pulmonary artery carries the blood to the lungs... where it picks up a fresh supply of oxygen...

This participant began by reading the introduction, which explains the basic functions of the circulatory system (segments 1 and 2). After reading a section of the introduction, he decided to take notes on this information (segment 3). The participant then continued this pattern of reading and taking notes. In segments 4 and 6, he read sections of the introduction, and then followed this reading by taking more notes on the information (segments 5 and 7). In segment 8, the participant switched from the introduction of the circulatory system to the heart article. As with the circulatory system, he began by reading a section of the article (segment 8) followed by the use of a strategy. In segments 9 and 11, he decided to re-read a section of the heart article and in segment 10 he summarized a section of the heart article.

These two coded transcriptions demonstrate how participants with various levels of prior domain knowledge self-regulate their learning differently, even when reading the same two sections in the hypermedia environment (the circulatory system article and the heart article). The participant who had high prior domain knowledge primarily relied on monitoring processes. An examination of the coded transcription indicates how this participant monitored her learning. The monitoring activities, which included monitoring both her understanding of the material and the relevance of specific sections in the environment, came immediately following brief examinations of the material in the environment. However, she did not exclusively use monitoring processes; she also intermittently summarized information she read. On the other hand, the participant who had low prior domain knowledge exclusively used strategies while reading sections from these two articles in the hypermedia environment. He primarily used the strategy of note-taking, but also summarized and re-read. Furthermore, an examination of this participant’s coded transcription suggests that he used strategies after reading a larger section of the text when compared to the participant with higher prior domain knowledge.

The third and fourth transcriptions come from different participants (one of whom demonstrated low prior domain knowledge while the other demonstrated high prior domain knowledge). These transcriptions highlight how participants with varying levels of prior domain knowledge use planning processes differently during learning. In the following examples, each participant is beginning the 40-min learning task and both start with the circulatory system article. This first transcription is from a participant whose pre-test indicated that he had high prior domain knowledge of the circulatory system.

<table>
<thead>
<tr>
<th>Segment</th>
<th>[SRL Process]SRL Category</th>
<th>[Sub-goal]Planning</th>
</tr>
</thead>
</table>
| 1       | What are the most important things the circulatory system does, I am going to look that up | [Planning]
| 2       | ...and I am going to go to introduction...and I am going to read introduction | [No Code] |
Reads: Circulatory System, or cardiovascular system, in humans, the combined function of the heart, blood, and blood vessels to transport oxygen and nutrients to organs and tissues throughout the body and carry away waste products.

Um...so, I would think that is the most important thing it does.

I am going to go on an article, to the components of the circulatory system.

Reads: The heart, blood, and blood vessels are the three structural elements that make up the circulatory system. The heart is the engine of the circulatory system. It is divided into four chambers: the right atrium, the right ventricle, the left atrium, and the left ventricle.

That makes sense.

I am going to click on that little guy...

I am going to click on heart so that I can read about the different parts of the heart.

This participant began the 40-min learning task (segment 1) by setting a sub-goal (i.e. identifying the most important things the circulatory system does). Based on this sub-goal, the participant went to the introduction of the circulatory system article (segment 2), and then read the first introductory paragraph (segment 3). After reading material that was relevant to his sub-goal, the participant made an inference (segment 4), and then continued with the circulatory system article (segment 5). After reading about the heart in the following section (segment 6), he monitored his understanding in segment 7 (“That makes sense”). He then decided to click on a hyperlink (segment 8), at which point he set another sub-goal of learning about the different parts of the heart. The next coded transcription is from a different participant whose pre-test indicated that she had low prior domain knowledge of the circulatory system.

Reads: Circulatory system, or cardiovascular system, in humans, the combined function of the heart, blood, and blood vessels to transport oxygen and nutrients to organs and tissues throughout the body and carry away waste products. Among its vital functions, the circulatory system increases the flow of blood to meet increased energy demands during exercise and regulates body temperature. In addition, when foreign substances or organisms invade the body, the circulatory system swiftly conveys disease-fighting elements of the immune system, such as white blood cells and antibodies, to regions under attack. Also, in the case of injury or bleeding, the circulatory system sends clotting cells and proteins to the affected site, which quickly stop bleeding and promote healing.
I am going to take notes... 

Reads: Among its vital functions, the circulatory system increases the flow of blood to meet increased energy demands during exercise and regulates body temperature... 

Taking notes... 

Reads: ...During each heartbeat, typically about 60–90 ml (about 2–3 oz) of blood are pumped out of the heart... 

Taking notes... 

As with the earlier participant, this participant began the learning task by reading the first introductory section in the circulatory system article (segments 1 and 2). However, no plan is made for the 40-min learning task. Instead, she read a substantial portion of the first section in this article, decided to take notes (segment 4), and then continued to read this section (segment 4). She continued with this pattern as she decided to take notes in segment 5, read the same section in segment 6, and then again decided to take some more notes in segment 7.

While the coded transcriptions from the first two participants provide a sense of how students with varying levels of prior domain knowledge use monitoring processes and strategies during learning, the transcriptions from the third and fourth participant highlight the differences in terms of planning processes. Both the third and fourth participant began the 40-min learning task with the introductory section of the circulatory system article. However, the participant with high prior domain knowledge initially set a sub-goal that was relevant to the overall goal of the learning task. After setting this sub-goal, he used various processes, including strategies (i.e. making an inference) and monitoring processes (i.e. expressing a feeling of knowing), until he once again set another relevant sub-goal (i.e. to learn about the different parts of the heart). On the other hand, the participant with low prior domain knowledge did not set a sub-goal when she started the learning task. Consistent with the quantitative results, this low prior domain knowledge participant primarily used strategies after reading substantial portions of text. Furthermore, she approached the learning task systematically by reading linearly and occasionally taking notes. There is no outward indication that a sub-goal, or any other planning processes, was guiding her learning.

8. Discussion

Previous research using various theoretical approaches, including the MDL (Murphy & Alexander, 2002) and CI (Kintsch, 1998) models have examined the relationship between prior domain knowledge and the learning process. For example, research using Kintsch’s CI model (1998) has demonstrated that students with domain-specific prior knowledge tend to use more advanced strategies, such as making inferences, when learning about science-related topics (McNamara et al., 1996). Similarly, research using the MDL model has found that use of more advanced strategies is related to the level of prior domain knowledge. A third theoretical approach, IPT (Winne, 2001; Winne & Hadwin, 1998), offers a framework that identifies the relationship between prior domain knowledge and strategies, as well as other processes that have been shown to be critical in learning, including monitoring and planning.
However, the advancement of this theoretical framework calls for empirical research that illustrates these relationships, especially in the emerging areas of research in learning with hypermedia (Azevedo, 2005). This study provides empirical and theoretically-driven data that extend the theoretical understanding of the relationship between prior domain knowledge and processes related to strategy use, monitoring, and planning.

Empirical results from the first research question indicated that prior domain knowledge was significantly related to the use of specific self-regulatory processes during learning. The results indicated that participants with higher prior domain knowledge tended to use significantly more planning and monitoring processes and fewer strategies when regulating their learning with hypermedia. This finding provides empirical evidence that supports the recent suggestion that prior domain knowledge may be necessary to monitor learning with hypermedia (Chen et al., 2006; Shapiro, 2004). The results also indicate that participants with lower prior domain knowledge tended to use more strategies when regulating their learning with hypermedia. However, these students tended to rely on a few, specific strategies during learning, such as note-taking and summarizing (see Table 5). They relied much less frequently on other strategies, such as making inferences, during learning with hypermedia. This finding is consistent with some research that has examined the relationship between prior domain knowledge and the use of strategies. For example, research in the field of text comprehension has found that students with lower prior domain knowledge are less likely to make inferences, especially when learning with low-coherence text (Best et al., 2005; McNamara, 2001). Similarly, research using the MDL indicates that students with lower prior domain knowledge tend to use less varied strategies during learning (Alexander et al., 1995; Murphy & Alexander, 2002).

Based on these findings, we draw several conclusions from the data. First, we argue that high prior domain knowledge learners already have a well-established, interconnected knowledge base of the topic which allows them to engage in what we call “knowledge verification”—i.e., they regulate their learning by using planning processes to activate their superior prior domain knowledge and then monitor their knowledge of the topic by comparing their current knowledge state with information provided in the hypermedia environment. When verifying their knowledge, students with higher prior domain knowledge may only use a strategy when their monitoring reveals a discrepancy between their current knowledge state and information presented in the environment. The coded transcripts presented in the discussion of research question two provide support for this conclusion. In the first coded transcription, the participant who had high prior domain knowledge demonstrated a pattern of reading a small chunk of the text in the hypermedia environment followed by a monitoring activity, such as feeling of knowing. This process of learning demonstrated the participant’s desire to verify her existing knowledge; that is, she repeatedly compared her conceptual knowledge structure of the circulatory system with the information provided in the text. When she acknowledged some overlap between her understanding and the information in the hypermedia environment (i.e. “I’m kind of remembering some of this stuff from bio in high school”), she moved to another section in the environment. As such, this participant was using her existing knowledge structure to guide her interaction with the hypermedia environment. However, when monitoring revealed a discrepancy between her understanding and the information in the environment, she used a strategy (i.e. summarization) to acquire this new knowledge.

Additionally, we conclude that in comparison to the high prior domain knowledge students, low prior domain knowledge students are engaging in what we call “knowledge
acquisition”—i.e., given their lack of a well-established knowledge base of the topic, they regulate their learning by frequently using a small subset of strategies (namely note-taking and summarizing) to learn as much as possible about the topic given the overall learning goal and the time limitations of the task. According to IPT, using other self-regulatory processes (such as planning and monitoring) consume valuable working memory space that could otherwise be used to process information during knowledge acquisition (Winne, 2001). As such, we conclude that the participants with low prior domain knowledge rarely planned or monitored their learning because using such processes would limit the available working memory capacity. The coded transcriptions presented in the discussion of research question two illustrate this relationship. In the second example, the participant, who had low prior domain knowledge of the circulatory system, read much larger chunks of the text, and typically followed this reading with a strategy such as note-taking or summarizing. Because this participant had low prior domain knowledge, he did not have a readily accessible knowledge structure that would have allowed him to monitor the relevancy of the content. Instead, he was in the initial stages of knowledge acquisition, and thus needed to use strategies to build a knowledge base of the circulatory system. The conclusions drawn from the data of this study provide theoretical and empirical implications.

8.1. Theoretical and educational implications

The theoretical implications from this study can be applied to educational practices. As suggested by IPT, students have a limited working memory capacity and thus are able to process a limited number of chunks of information. When students are faced with academic tasks in which they have very little prior domain knowledge, they may not be able to simultaneously process information and apply self-regulatory processes (Winne, 2001). In order to process relatively novel information, a substantial portion of working memory may be consumed, leaving little or insufficient working memory capacity for the use of self-regulatory processes. If, however, students with lower prior domain knowledge attempt to use a variety of self-regulatory processes, they then may have limited, if insufficient, working memory capacity to process information (Kanfer & Ackerman, 1989). As such, students with low prior domain knowledge may be caught in a catch 22. If they focus on self-regulating their learning, they may not be able to process information. If, however, they decide to focus primarily on processing information, they may not be able to simultaneously self-regulate their learning.

In order to address this potentially problematic issue, educators should carefully consider the self-regulatory demands of learning contexts, especially hypermedia environments where students are faced with complicated tasks in which they have little prior domain knowledge. In these environments, educators should structure the environment in a manner that limits the need for students to simultaneously self-regulate their learning as they process information. Additionally, educators can facilitate the use of several strategies that may help in maximizing the availability of working memory capacity for processing information. For example, strategies such as note-taking allow for the external recoding of information so that working memory is not overburdened. Educators should explicitly promote the use of note-taking during complicated tasks because this may allow students to free up cognitive resources for processing information.

Additionally, educators should externally support the use of specific SRL processes that have been shown to be related to learning, especially when students are learning about a
domain in which they have little prior domain knowledge. For example, students with little prior domain knowledge who attempt to use metacognitive monitoring processes (such as JOL and FOK) may have little working memory capacity to allocate to processing information. Thus, educators should structure the environment so that there are limited demands on students’ working memory in terms of metacognitive monitoring. For example, educators’ detailed feedback, as a by-product of external monitoring, may alleviate students’ need to monitor their own progress, which in turn maximizes working memory space.

8.2. Future directions

While this study contributes to our understanding of the relationship between prior domain knowledge and how undergraduates self-regulate their learning with hypermedia, future research could extend this research agenda by addressing the following two issues. First, this research agenda calls for a more comprehensive empirical examination of learning with hypermedia. As previous research has demonstrated, there certainly are factors other than prior domain knowledge which may explain students’ use of specific self-regulatory processes during learning. For example, previous research has demonstrated a relationship between various motivational constructs and self-regulated learning with hypermedia (e.g., goal orientation; Moos & Azevedo, 2006). As such, examining the relationship between theoretically-grounded constructs of motivation, prior domain knowledge, and SRL will provide a more comprehensive model of learning with hypermedia (Winne, 2005). In addition to motivational constructs, future research should consider examining the relationship between cognitive load, prior domain knowledge, and SRL. Though this study did not empirically examine cognitive load, previous research has indicated that multiple representations of information, as found in hypermedia, may present redundant material (Sweller, 2005). This redundancy increases working memory load, which can impede learning (Leahy, Chandler, & Sweller, 2003). This line of research has indicated that cognitive load is a valid issue to study in the area of learning with hypermedia, and thus future research should consider the relationship between cognitive load and prior domain knowledge.

Additionally, though many strengths have been identified with the IPT approach to SRL (Zimmerman & Schunk, 2001), there are some theoretical concerns that future research should address. Namely, this theoretical approach does not currently account for individual differences in how students respond to negative discrepancies in unfamiliar and dynamic environments (Zimmerman & Schunk, 2001). For example, when students monitor their learning and identify discrepancies between their goal(s) and current knowledge state, some students adapt their use of SRL processes and persevere while other students lower their standards and do not change their use of SRL processes. In order to extend the IPT approach to SRL, theoreticians and researchers should address how this theoretical framework explains variability in students’ SRL when they face negative feedback.

8.3. Limitations

There are several limitations of this study that need to be addressed. First, the relationship between prior domain knowledge and use of SRL processes, as found in this study, may only hold true for college students. Previous research indicates that developmental constraints on SRL exist and that younger students (especially those in elementary and
middle school) may have limited capacity to self-regulate their learning (Pintrich & Zusho, 2002). Thus, the relationship between prior domain knowledge and use of specific SRL processes during learning with hypermedia, as demonstrated in this study, may differ with younger students. Furthermore, previous research has demonstrated that prior experience with specific types of CBLEs is strongly related to learning with these environments (Hasan, 2003; Salanova, Grau, & Cifre, 2000). While undergraduates may have higher prior domain knowledge of the circulatory system, it is quite possible that younger students may have more experience learning with hypermedia environments. However, hypermedia experience was not measured in this study, and thus the external validity of this study is somewhat limited. In order to address this concern, future research should consider empirically examining prior domain knowledge and prior hypermedia experience for both undergraduates and younger learners. Secondly, gender differences may exist in the development of SRL processes, and these differences may potentially moderate the relationship between prior domain knowledge and use of SRL processes (Pintrich & Zusho, 2002). However, the original research questions for this study did not focus on gender differences and thus the potential moderating effect of this factor was not reported. Future research should consider addressing the potential moderating effects of gender on the relationship between prior domain knowledge and self-regulated learning with hypermedia.

Appendix A

Classes, descriptions and examples of variables used to code students’ self-regulatory behavior (based on Azevedo, Cromley, et al., 2004)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Student Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning</td>
<td>A plan involves coordinating the selection of operators. Its execution involves making behavior conditional on the state of the problem and a hierarchy of goals and sub-goals</td>
<td>“First I’ll look around to see the structure of environment and then I’ll go to specific sections of the circulatory system”</td>
</tr>
<tr>
<td>Sub-Goals</td>
<td>Consist either of operations that are possible, postponed, or intended, or of states that are expected to be obtained. Goals can be identified because they have no reference to already-existing states</td>
<td>“I’m looking for something that’s going to discuss how things move through the system”</td>
</tr>
<tr>
<td>Prior domain knowledge</td>
<td>Searching memory for relevant prior domain knowledge either before beginning of a task or during task performance</td>
<td>“It’s hard for me to understand, but I vaguely remember learning about the role of blood in high school”</td>
</tr>
<tr>
<td>Activation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recycle Goal in Working Memory</td>
<td>Restating the goal (e.g., question or parts of a question) in working memory</td>
<td>“Make sure you learn about the different parts of the heart and their purpose...okay...”</td>
</tr>
<tr>
<td>Monitoring</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Judgment of Learning</td>
<td>Learner becomes aware that they do not know or understand everything they read</td>
<td>“I don’t know this stuff, it’s difficult for me”</td>
</tr>
<tr>
<td>Feeling of Knowing</td>
<td>Learner is aware of having read something in the past and having some understanding</td>
<td>“...let me read this again since I’m starting to get it...”</td>
</tr>
</tbody>
</table>
### Appendix A (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description*</th>
<th>Student Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Questioning</td>
<td>Posing a question and re-reading to improve understanding of the content</td>
<td>[Learner spends time reading text] and then states “what do I know from this?” and reviews the same content</td>
</tr>
<tr>
<td>Content Evaluation</td>
<td>Monitoring content relative to goals</td>
<td>“I’m reading through the info but it’s not specific enough for what I’m looking for”</td>
</tr>
<tr>
<td>Identify Adequacy of Information</td>
<td>Assessing the usefulness and/or adequacy of the content (reading, watching, etc.)</td>
<td>“…structures of the heart…here we go…”</td>
</tr>
<tr>
<td>Monitor Progress Toward Goals</td>
<td>Assessing whether previously-set goal has been met.</td>
<td>“Those were our goals, we accomplished them”</td>
</tr>
<tr>
<td>Monitor Use of Strategies</td>
<td>Participant comments on how useful a strategy was</td>
<td>“Yeah, drawing it really helped me understand how blood flow throughout the heart”</td>
</tr>
<tr>
<td>Strategy use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selecting a New Informational Source</td>
<td>The selection and use of various cognitive strategies for memory, learning, reasoning, problem solving, and thinking</td>
<td>[Learner reads about location valves] then switches to watching the video to see their location</td>
</tr>
<tr>
<td>Coordinating Informational Sources</td>
<td>Coordinating multiple representations, e.g., drawing and notes</td>
<td>“I’m going to put that [text] with the diagram”</td>
</tr>
<tr>
<td>Read New Paragraph</td>
<td>The selection and use of a new paragraph</td>
<td>“OK, now on to pulmonary”</td>
</tr>
<tr>
<td>Review Notes</td>
<td>Reviewing learner’s notes</td>
<td>“Carry blood away. Arteries—away.”</td>
</tr>
<tr>
<td>Memorization</td>
<td>Learner tries to memorize text, diagram, etc.</td>
<td>“I’m going to try to memorize this picture”</td>
</tr>
<tr>
<td>Free Search</td>
<td>Searching the hypermedia environment without specifying a specific plan or goal</td>
<td>“I’m going to the top of the page to see what is there”</td>
</tr>
<tr>
<td>Goal-Directed Search</td>
<td>Searching the hypermedia environment after specifying a specific plan or goal</td>
<td>Learner types in blood circulation in the search feature</td>
</tr>
<tr>
<td>Summarization</td>
<td>Summarizing what was just read, inspected, or heard in the hypermedia environment</td>
<td>“This says that white blood cells are involved in destroying foreign bodies”</td>
</tr>
<tr>
<td>Taking Notes</td>
<td>Copying text from the hypermedia environment</td>
<td>“I’m going to write that under heart”</td>
</tr>
<tr>
<td>Draw</td>
<td>Making a drawing or diagram to assist in learning</td>
<td>“…I’m trying to draw the diagram as best as possible”</td>
</tr>
<tr>
<td>Re-reading</td>
<td>Re-reading or revisiting a section of the hypermedia environment</td>
<td>“I’m reading this again.”</td>
</tr>
<tr>
<td>Inferences</td>
<td>Making inferences based on what was read, seen, or heard in the hypermedia environment</td>
<td>…[Learner sees the diagram of the heart] and states “so the blood…through the…then goes from the atrium to the ventricle…and then…”</td>
</tr>
<tr>
<td>Hypothesizing</td>
<td>Asking questions that go beyond what was read, seen or heard</td>
<td>“I wonder why just having smooth walls in the vessels prevent blood clots from forming…I wish they explained that…”</td>
</tr>
<tr>
<td>Knowledge Elaboration</td>
<td>Elaborating on what was just read, seen, or heard with prior domain knowledge</td>
<td>[after inspecting a picture of the major valves of the heart] the learner states “so that’s how the systemic and pulmonary systems work together”</td>
</tr>
</tbody>
</table>

(continued on next page)
Appendix A (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Student Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mnemonic</td>
<td>Using a verbal or visual memory technique to remember content</td>
<td>“Arteries—A for away”</td>
</tr>
<tr>
<td>Evaluate Content as Answer to Goal</td>
<td>Statement that what was just read and/or seen meets a goal or sub-goal</td>
<td>[Learner reads text]. . .” So, I think that’s the answer to this question”</td>
</tr>
</tbody>
</table>

* All codes refer to what was recorded in the verbal protocols (i.e., read, seen, or heard in the environment and/or during discussions).

Appendix B. One page from a coded transcript of a high prior domain knowledge student

**Condition B/Study 4/ UG114 “NP”**

11/18/03

important. In the systemic circulation The heart ejects oxygen-rich blood under high pressure out of the heart’s main pumping chamber, the left ventricle that’s probably important main pump chamber is the left ventricle and the largest artery is the aorta which I am also going to write down...largest artery aorta, um these small arteries branch off from it and then they branch out into even smaller ones called arterioles and they become smaller and become capillaries so it’s kind of like a chain, and then once the blood reaches the capillaries, blood pressure is reduced. Capillaries have very thin walls that permit dissolved oxygen and nutrients from the blood to diffuse across a fluid, that fills the gaps between the cells of tissues or organs. dissolved oxygen and nutrients then enter the cells from the interstitial fluid by diffusion also know that from bio uh, Meanwhile, carbon dioxide and other wastes leave the cell, diffuse through the interstitial fluid, cross capillary walls, and enter the blood. In this way, the blood delivers nutrients and removes wastes without leaving the capillary tube. Um After delivering oxygen the deoxygenated blood in the capillaries then goes back to the heart. The capillaries merge to form venules. I’m going to write that down capillaries form venules and These veins join to get larger and then, the veins converge into two large veins: the inferior vena cava, which is the lower half; and the superior vena cava, which brings blood from the upper half. And both of these two large veins join at the right atrium so I’ll just put cava’s join at right atrium um now that there’s no
References


