CAN THE USE OF COGNITIVE AND METACOGNITIVE SELF-REGULATED LEARNING STRATEGIES BE PREDICTED BY LEARNERS' LEVELS OF PRIOR KNOWLEDGE IN HYPERMEDIA-LEARNING ENVIRONMENTS?

Submitted by: Michelle Taub ©

Department of Educational and Counseling Psychology

McGill University, Montreal

July 2013 ©

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Arts, Educational Psychology, Learning Sciences Stream

Table of Contents

Abstract

Research on self-regulated learning (SRL) in hypermedia-learning environments is a growing area of interest. One major issue is the role of prior knowledge, which can influence how students interact with hypermedia-learning systems in terms of their use of SRL strategies. In this study, we investigated 52 undergraduate participants' interactions with MetaTutor, a multi-agent, hypermedia-based learning environment, which teaches students about the human circulatory system. We assessed how students' prior knowledge levels may have affected how they used particular cognitive and metacognitive SRL strategies in terms of frequency of use; most common sequences of use; the number engaged in; time spent engaging; and relevant pages visited. We expected that overall, students with high prior knowledge would engage in significantly more cognitive and metacognitive SRL strategies than students with low prior knowledge. Moreover, we expected that students with different prior knowledge levels would engage in different sequences of SRL strategies. We also predicted that students with high prior knowledge would engage in larger amounts of SRL processes, spend more time engaging in SRL processes, and visit more pages, which were relevant to their current sub goals than students with low prior knowledge. Results showed significant differences in the total use of SRL strategies between high and low prior knowledge groups. More specifically, results revealed significant differences in the total use of metacognitive strategies, but not total cognitive strategies between prior knowledge groups, which supported the hypothesis that states students with high prior knowledge would engage in significantly more metacognitive SRL strategies that students with low prior knowledge, but did not support hypothesis, which claimed that students with high prior knowledge would engage in significantly more cognitive strategies than students with low prior knowledge. Results also revealed different sequences of use of SRL strategies between prior knowledge groups; and showed that students used different numbers of SRL processes and thus spent different times engaging in SRL processes; although students with both levels of prior knowledge did not seem to visit different numbers of relevant pages. These results, therefore, supported the hypothesis that students in different prior knowledge groups would engage in different sequences of SRL strategies. Furthermore, these results reveal that the amount of SRL strategies students engaged in, the time students spent engaging in SRL processes, and the relevant pages students visited depended on the sub goal they were working on. These results can have important implications on how we design multi-agent, hypermedia environments, such that we can design pedagogical agents that can adapt to students' learning needs, based on their prior knowledge of the content. Future studies should investigate student learning using multi-channeled data, which can help researchers design agent-based learning environments that adapt to individual needs in terms of cognitive, metacognitive, motivational, and affective processes.

Résumé

L’étude de l’apprentissage autorégulé (AAR) dans des environnements hypermédias d’apprentissage est un domaine de recherche en développement. Une question essentielle concerne le rôle des connaissances antérieures, qui peuvent influencer la manière dont les étudiants utilisent des stratégies d’AAR lorsqu’ils interagissent avec des systèmes hypermédias d’apprentissage. Dans cette étude, nous examinons les interactions de 52 étudiants de premier cycle universitaire avec MetaTutor, un environnement d’apprentissage multi-agents à base d’hypermédia portant sur le système circulatoire sanguin. Nous évaluons comment des niveaux de connaissances antérieures différents ont pu affecter l’usage par les étudiants de certaines stratégies cognitives et métacognitives d’AAR en termes de fréquence d’utilisation, de séquences d’actions les plus couramment utilisées, du nombre de fois que ces séquences ont été utilisées, du temps passé sur celles-ci et du nombre de pages visitées pertinentes pour leurs sous-buts. De manière générale, nous nous attendions à ce que les étudiants avec des connaissances antérieures importantes utilisent davantage de stratégies cognitives et métacognitives d’AAR que les étudiants ayant peu de connaissances antérieures. De plus, nous nous attendions à voir les étudiants avec différents niveaux de connaissances antérieures employer des séquences de stratégies d’AAR différentes. Nous avions également prédit que les étudiants avec des connaissances antérieures importantes tendraient à employer davantage de processus d’AAR, à passer plus de temps à effectuer ces processus, et à visiter plus de pages pertinentes pour leurs sous-buts actuels que les étudiants avec de faibles connaissances antérieures. Les résultats montrent l’existence de différences significatives dans l’usage total des stratégies d’AAR entre le groupe ayant de faibles connaissances antérieures et

celui ayant un niveau élevé de connaissances antérieures. En particulier, les résultats révèlent des différences significatives dans l'usage total des stratégies métacognitives, mais pas dans l'usage total des stratégies cognitives entre les deux groupes, ce qui confirme l'hypothèse selon laquelle les étudiants avec des connaissances antérieures importantes emploieraient significativement plus de stratégies métacognitives d'AAR que les étudiants avec de faibles connaissances antérieures, mais ne valide pas l'hypothèse selon laquelle les étudiants avec des connaissances antérieures importantes emploieraient plus de stratégies cognitives que les étudiants avec de faibles connaissances antérieures. Les résultats montrent également différentes séquences d'utilisation de stratégies d'AAR entre les deux groupes ayant des niveaux différents de connaissances antérieures : les étudiants ont utilisé un nombre différent de processus d'AAR et ont ainsi passé des temps différents à déployer ces processus ; en revanche, les différences de niveaux de connaissances antérieures n'ont pas semblé avoir d'influence sur le nombre de pages pertinentes visitées. Ces résultats valident ainsi l'hypothèse selon laquelle des étudiants avec différents niveaux de connaissances antérieures emploieraient différentes séquences de stratégies d'AAR. En outre, ces résultats révèlent que le nombre stratégies d'AAR employé par les étudiants, le temps passé à déployer des processus d'AAR, et les pages pertinentes visitées par les étudiants dépendent du sous-but sur lequel ils travaillent. Ces résultats peuvent avoir des implications importantes pour la conception d'environnements hypermédias multi-agents, de manière à pouvoir concevoir des agents pédagogiques capables de s'adapter aux besoins d'apprentissage des étudiants, sur la base de leurs connaissances antérieures du contenu. Des études ultérieures devront examiner l'utilisation de données multi-canaux, qui peuvent aider les chercheurs à concevoir des environnements d’apprentissage à base d’agents capables de s’adapter aux besoins individuels en termes de processus cognitifs, métacognitifs, motivationnels et affectifs.

Acknowledgements

There are many people that I would like to thank in assisting me to complete my Master's Thesis. First, I'd like to thank my supervisor, Dr. Roger Azevedo, for I would not have been able to write this Thesis without his guidance, support, and feedback. In addition, Dr. Azevedo has allowed me to use data from his MetaTutor Grant in performing the analyses that I did for this study. Dr. Azevedo has been an excellent mentor and supervisor throughout the past two years, and I could not have accomplished anything I have accomplished without him. I am so grateful to have had such an ambitious and helpful supervisor, and I look forward to pursuing my PhD under your supervision. Thank you for everything!

I would also like to thank my fellow post-doctoral and graduate research assistants, who have assisted and supported me throughout the past two years. Thank you François Bouchet, Babak Khosravifar and Dean Bhageerutty for all your hard work in extracting the data for this Thesis. You never said no to any of my ideas for data extraction, and this Thesis would not be possible without all your hard work. Thank you Nicole Pacampara for sharing your illustrative talents and helping me in creating the Figures presented in this Thesis. To Melissa Duffy, Reza Feyzi-Behnagh, Jason Harley, and Gregory Trevors, thank you for assisting and guiding me throughout obtaining my Master's. I have enjoyed working with you on projects and collaborating on papers, and I look forward to pursuing many more publications together throughout all of our academic careers. I have truly enjoyed the time I have spent with all of you, and these two years would not have been the same without any of you.

To all of the other research assistants, thank you for all your hard work in the lab. You have all contributed to my positive research experience in the S.M.A.R.T. Lab, and I am so glad to have met and worked with all of you. Thank you to Lauren Agnew, Kelsey Anderson, Valérie Bélanger-Cantara, Sophie Griscom, Lana Karabachian, Faiz Khan, Nicholas Mudrick, Aly Segura, Victoria Stead, Jack Tokarz, and Wook Yang for everything you worked on in the lab to make data collection and analysis run that much smoother. I would not have been able to complete my thesis without your help in data collection with MetaTutor, and I appreciate everything you have done.

Finally, I want to thank my family and friends for supporting me throughout the past two years.

Support for this project was made possible by funding for the National Science Foundation, Social Sciences and Humanities Research Council of Canada, and the Canada Research Chair awarded to Dr. Azevedo.

Introduction

Self-regulated learning, SRL, is an important educational construct, which has been shown to be effective for students as they learn and study various subjects (Azevedo, 2005, 2007; Winne & Perry, 2000; Zimmerman & Schunk, 2011). When students self-regulate their learning, they are playing an active role in the learning process by engaging in planning, goal-setting, and other cognitive and metacognitive processes (Azevedo, 2005). Research has found that when students engage in self-regulated learning, they achieve high learning outcomes (Greene & Azevedo, 2010; Azevedo et al., 2013). It is, therefore, important for students to engage in self-regulated learning skills, such as planning or monitoring, and strategies, such as judgment of learning or summarizing, in order to achieve and learn to their greatest capacity.

Despite the increasing evidence for the effectiveness of self-regulated learning for students' learning outcomes (Azevedo et al., 2010, 2012), research has also revealed that students do not enact these effective self-regulated learning strategies during learning (Azevedo, 2005; Azevedo et al., 2012). Interdisciplinary researchers have been designing and developing computer-based learning environments (CBLEs) (e.g., multimedia, hypermedia, intelligent tutoring systems, multi-agent systems) to foster and promote effective self-regulated learning in students as they learn about various topics, such as biology, physics, and ecology (Aleven et al., 2010; Azevedo & Aleven, 2013; Azevedo et al., 2013; Biswas et al., 2010; Graesser et al., 2007; Jonassen & Land, 2012; D'Mello et al., 2013; Lajoie et al., 2013; Lester et al., 2013; Woolf, 2009).

Some CBLEs are multi-agent, such that they are programmed to include pedagogical agents, PAs, who are present to assist students by providing scaffolding and feedback during learning, problem solving, strategy training, and skill acquisition (Azevedo et al., 2012a; Biswas et al., 2010; D'Mello et al., 2013; Graesser & McNamara, 2010; Lester et al., 2013). In addition to content learning, these agents are programmed to assist learning about different aspects of SRL, such as planning, goal-setting, metacognitive monitoring, strategy use, and reflection (see Azevedo & Aleven, 2013; Azevedo et al, 2012). The use of PAs can be effective for learners because research has shown that when students are provided with the appropriate scaffolding, this can help them to better learn (Kinnebrew et al., 2013), and more specifically to self-regulate their learning (Azevedo et al., 2012a; Graesser & McNamara, 2010; Winters et al., 2008). The role of prior knowledge is a critical individual differences variable that has not been adequately examined in the context of SRL and learning with multi-agent systems. Results will contribute to theoretical (e.g., to understand the deployment of SRL processes, based on prior knowledge) and educational (e.g., to provide the dynamic assessment and differential scaffolding, based on learners' prior knowledge) implications to SRL, which can assist researchers in designing CBLEs that adapt to student characteristics, such as level of prior knowledge.

The focus of this study is to assess how students' prior knowledge can impact the way they self-regulate their learning in a CBLE, with the assistance of pedagogical agents. Prior knowledge of the domain can greatly affect how students engage in different SRL processes and use learning strategies (Shapiro, 2004; Moos & Azevedo, 2008); and so when we are creating these environments, it is important to consider how students' prior knowledge of the domain can potentially influence the SRL skills (e.g., planning and monitoring) and strategies, which can be metacognitive or cognitive (e.g., content evaluation and note-taking) that they use. For this study, we acknowledged past findings regarding the importance of prior knowledge, and thus assessed how prior knowledge of the circulatory system influenced how students used cognitive and metacognitive SRL strategies as they learned with MetaTutor, a multi-agent, adaptive hypermedia-learning environment (Azevedo et al., 2009, 2010, 2011, 2012, 2013).

Theoretical Framework

In our analyses of self-regulated learning, we view SRL as an event, which temporally unfolds in real-time (Winne & Perry, 2000). We used Winne and Hadwin's Information-Processing Model (1998, 2008) as our theoretical model of SRL, according to which, learning occurs in four basic phases: (a) definition of the task, (b) setting goals and planning, (c) studying tactics, and (d) adaptations; and information processing occurs within each learning phase (Winne & Hadwin, 1998). In phase one, the learner assesses the task at hand, and determines the environmental factors that are available in helping to learn, and thus to accomplish the task. Phase two involves planning and goal setting, in which the learner is to set the goals, which are needed in order to accomplish the task. In addition, in this phase, the learner plans the appropriate sub-tasks needed in order to complete the sub-goals that were set at the beginning of the phase. Phase three involves the learner engaging in the strategies that he or she planned to engage in during the second phase. In addition, the learner monitors the progress he or she is making towards achieving the goals that have been set. Lastly, phase four is characterized by a reflection of what was accomplished in phases one through three. The learner makes the appropriate adaptions to plans and goals that were set, which can be based on the learner's modifications to his or her understanding of the task (Winne & Hadwin, 1998).

It is expected that students with high prior domain knowledge would progress through each stage differently than students with low prior knowledge of the domain, for which the task focuses on. In phase one (defining the task), students with high prior knowledge would not differ from students with low prior knowledge in terms of defining what the task is asking of them; however students with high prior knowledge will be more aware of the environmental factors, which can be used as resources in accomplishing the task, compared to students with low prior knowledge, who will have difficulty identifying the appropriate environmental factors, which can be used in solving the task. For example, students with high prior knowledge might recall how they approached a similar problem in the past, however students with low prior knowledge will not make this connection between the present and past problems. In the second phase (setting goals and planning), students with high prior knowledge will not have difficulties planning or creating the sub-goals necessary to achieve the task; however since students with low prior knowledge are not familiar with the domain, they will experience difficulties in creating sub-goals needed in order to accomplish the task. For example, students with high prior knowledge know that if they are learning about the circulatory system, they will need to create a sub-goal that either relates to prior knowledge or goes beyond their prior knowledge (of the circulatory system) and deals specifically with the overall learning goal set by the researchers. Students with low prior knowledge, however, may experience difficulties creating and prioritizing relevant sub-goals, given their lack of domain knowledge. For phase three (studying tactics), both prior knowledge groups will be able to deploy the strategies that they have set to do; however students in the high prior knowledge group will differ by deploying more sophisticated and effective use of these strategies, compared to students with low prior knowledge. For example, students in both groups could have planned to take notes, but students with high prior knowledge may translate the notes into their own words, whereas students with low prior knowledge may copy the words verbatim from the text. In addition, high prior knowledge students will be able to metacognitively monitor their emerging understanding of the topic more accurately than those with low prior knowledge. Finally, in phase four, students with high prior knowledge will be able to reflect on their learning, and adjust their understanding of the question, whereas students with low prior knowledge may not be able to make such a reflection. For example, students with high prior knowledge may have planned to spend a particular amount of time achieving a sub-goal; however during reflection, the student may realize that he or she needs more time to accomplish another sub-goal. A student with low prior knowledge, however, may not be able to reflect on the time he or she has allotted to completing the sub-goals.

This study focuses on the role of students' prior knowledge on their use of self-regulated learning strategies, and whether or not we see significant differences between high and low prior knowledge groups. Winne and Hadwin's model (2008) emphasizes the role of prior knowledge as a key factor in self-regulated learning, however there is no specific framework or hypotheses, which have been generated, that address the role of prior knowledge in self-regulated learning with computer-based learning environments. Thus, the following study acts as a preliminary set of results, which can be used to facilitate the beginning of the development of a framework for the role of prior knowledge in self-regulated learning with hypermedia (Taub et al., 2013).

Based on existing literature on prior knowledge and Winne and Hadwin's SRL model, we make the following assumptions: (1) learners with high prior domain knowledge (HPK) will be more effective at self-regulating their learning, compared to students with low prior domain knowledge (LPK), because they have more relevant domain knowledge that allows them to anchor new knowledge to existing knowledge (Mayer, 2004); (2) Learners with high prior domain knowledge (HPK) will be more effective at self-regulating their learning, compared to students with low prior domain knowledge (LPK), because they have more memory space to engage in metacognitive monitoring and use sophisticated cognitive strategies, which can help them to better self-regulate their learning (Paas et al., 2013).

In MetaTutor, several cognitive and metacognitive processes are available for students to engage in as they learn about the circulatory system (Azevedo et al., 2012). Based on the abovementioned assumptions, several specific assumptions have been created for the following study, regarding the role of self-regulated learning processes with MetaTutor between students with high, compared to low, prior knowledge. These assumptions state that: (1) Overall, we expect that, according to the concept of cognitive load (Paas et al., 2013), students with high prior knowledge will engage in more cognitive and metacognitive processes combined during learning than students with low prior knowledge. (2) More specifically, we expect that students with high prior domain knowledge on the human circulatory system will engage in more cognitive processes; take notes (TN), make inferences (INF), create summaries (SUMM), and activate prior knowledge (PKA) than students with low prior domain knowledge of the human circulatory system. (3) We also expect that students with high prior knowledge will engage in more metacognitive processes; judgment of learning (JOL), feeling of knowing (FOK), content evaluation (CE), monitoring progress towards goals (MPTG), and monitoring, compared to students with low prior knowledge. Furthermore, based on data mining hypotheses; (4) we expect that students with high prior knowledge will engage in more effective sequences of use of SRL strategies, which implies that students will engage in metacognitive strategies prior to engaging in cognitive strategies. For example, it is more effective for one to judge whether or not he or she understand the content (i.e., does a JOL) before he or she takes notes (TN) on such content. Students in the low prior knowledge group, however, may not make this differentiation, and will engage in both cognitive and metacognitive strategies when they start a new page or sub-goal; (5) we expect that students with high prior knowledge will engage in more SRL strategies and spend more time engaging in SRL strategies as they accomplish the sub goal they are working on during learning, compared to students with low prior knowledge, who will spend more time reading the content and will engage in fewer SRL strategies and thus spend less time engaging in SRL strategies; and (6) we expect that students with high prior knowledge will visit more pages that are relevant to the sub goal they are working on, compared to students with low prior knowledge, who will visit more pages, and not just those that are relevant to their sub goal.

Literature Review: Prior Knowledge, SRL, and CBLEs

Prior knowledge is an important factor in learning, and it is important for researchers to investigate and understand its role in students' self-regulated learning. When students self-regulate, they engage in cognitive, affective, metacognitive, and motivational (CAMM) processes, which can all be influenced by their prior knowledge levels, and therefore their individual needs. Consequently, we must adapt our instructional methods in designing hypermedia-learning environments in order to cater to students' individual learning needs in order to promote effective use of CAMM processes to self-regulate their learning (Azevedo et al., 2012).

Previous research has examined the effect of prior knowledge on one's learning (Ericsson et al., 2006; Sawyer, 2006; Zimmerman & Schunk, 2011), however there is limited research investigating the role that prior knowledge plays in affecting how one self-regulates his or her learning with complex topics with hypermedia environments. Shapiro (2004) conducted two experiments based on previous research from Voss & Silfies (1996), which demonstrated that when students were presented with fictional text the authors created, students with high prior knowledge relied on accurate information they had previously known regarding the subject. Shapiro's first experiment tested whether domain knowledge played an important role in learning, even when the text is fictional (Shapiro, 2004). Since the content was fictional, topic knowledge was accounted for because participants could not rely on their previous knowledge of that specific topic; they could only rely on their previous knowledge of the domain. These experiments were thus testing for participants' prior domain knowledge and its impact on learning (Shapiro, 2004). Results demonstrated that domain knowledge significantly contributed to learning more than any other variable, which additionally demonstrates that fictional text cannot eliminate the effect of prior domain knowledge on learning (Shapiro, 2004). In her second experiment, Shapiro tested novices to prove that novices have high topic knowledge. The purpose of this experiment was additionally meant to demonstrate that the effect of prior knowledge occurs naturally, and not solely in a laboratory setting

(Shapiro, 2004). The results demonstrated that topic knowledge played a significant role in most of the post-test scores obtained. Results also confirmed that students with low prior knowledge show a large range of topic knowledge in a naturalistic setting (Shapiro, 2004). Shapiro was thus able to conclude that prior knowledge influences how students self-regulate their learning.

Although there is limited research on the role of prior knowledge in assessing students' self-regulated learning with hypermedia, some researchers have conducted studies, which have assessed prior knowledge in multimedia learning environments. Winters and Azevedo (2005) used GenScope, a computer-based learning environment, which teaches students about genetics. Students were paired into dyads, based on their prior knowledge, determined by pre-test data. Their results indicated that based on the pre- to post-test data shift, students with low prior knowledge displayed an increase in their understanding of the material, while students with high prior knowledge did not present any changes in their understanding of the content (Winters & Azevedo, 2005). Their analyses of students' verbalizations revealed that students with low prior knowledge relied on others to help facilitate their use of self-regulated learning compared to students with high prior knowledge, who, during the study, regulated their own learning and provided support for the low prior knowledge students who were seeking assistance (Winters & Azevedo, 2005). These results were important in demonstrating that the appropriate scaffolds can be provided to help foster the use of effective self-regulated learning strategies (Winters & Azevedo, 2005).

Moreover, Scheiter et al (2009) conducted a study, where they used cluster analysis and determined that there were five clusters of students based on their characteristics. The characteristics used in their study were strategies for the use of information, cognitive load, and learning outcomes in a hypermedia environment (Scheiter et al., 2009). Their results demonstrated that when students had characteristics, which were seen as more favorable, they showed better use of strategies for information-use, less use of cognitive load, and solved more problems correctly, compared to students with less favorable characteristics. Favorable characteristics included high prior knowledge, complex epistemological beliefs, positive attitudes towards math, and better use of cognitive and metacognitive strategies (Scheiter et al., 2009). High prior knowledge, therefore, played a positive factor in influencing student learning characteristics.

Moos and Azevedo (2008) conducted a study where they examined the relationship between prior knowledge and self-regulated learning with hypermedia. They implemented a 40-minute learning session, where the results demonstrated that prior knowledge is significantly related to students' use of self-regulated learning during learning with hypermedia (Moos & Azevedo, 2008). More specifically, the results revealed that prior knowledge was positively related to monitoring and planning, and was negatively related to strategies (Moos & Azevedo, 2008).

In a subsequent study, Moos and Azevedo (2009) implemented a 30-minute hypermedia task and used a series of science questions (Moos & Azevedo, 2009) to foster conceptual understanding about the human circulatory system. First, they found that self-efficacy and monitoring processes were significantly related; second, they found that prior domain knowledge and monitoring understanding were significantly related; and finally, they found that, through regression analysis, the relationship between self-efficacy and hypermedia learning was mediated by students' monitoring levels and by the environment (Moos & Azevedo, 2009).

Several studies have not found a significant effect of prior knowledge on learning outcomes. Shapiro (1999) used interactive overviews (IOs) on ecology in order to help students meet their learning goals. She based the study on Kintsch's construction integration model (1988), which explains that in order for deep learning to occur, one must integrate this newly learned information with his or her prior knowledge of the content. However, results demonstrated that IOs were helpful tools for students to achieve their learning goal, which they did not have any prior knowledge on (Shapiro, 1999). Van Seters et al. (2012) used e-learning materials to demonstrate how students work differently, based on their own characteristics. In order to determine these student characteristics, the authors collected participants' demographic information; they measured participants' motivation and prior knowledge levels, and their use of learning strategies and the learning paths they followed (van Seters et al., 2012). Results showed that students did follow different learning paths; and there were significant differences among Dutch BSc students and international MSc students in terms of their intrinsic motivation, the learning paths they followed, and the learning strategies they used. However, they found that prior knowledge did not have an effect on students' learning paths (van Seters et al., 2012).

The studies that have been conducted were influential for a number of reasons. The majority of studies are being conducted in naturalistic settings, which increases the generalizability of such studies. The more results we can find outside of the laboratory, the better able we will be able to apply our findings to other real classroom settings.

Laboratory studies are beneficial because they allow researchers to gain full control over the experimental setting, however they lack ecological validity, which limits our findings for implying such results to real-world settings. These studies also included the analysis of many factors, along with prior knowledge, and so do not limit the findings to one sole factor. We know that in real-world classroom settings, there are many factors that can influence learning, and so it is important to consider multiple factors when conducting research. For example, van Seters et al. (2012) included prior knowledge, motivation, and learning paths and strategies when observing student characteristics; and Scheiter et al. (2009) considered strategy-use, cognitive load, and learning outcomes in determining student characteristics. Moreover, Moos & Azevedo (2009) examined multiple relationships between different cognitive, metacognitive, and motivational aspects in hypermedia learning environments.

Additionally, I stressed the importance of generalizability of findings because many studies are not exact replicas of real-life classrooms. For example, Moos & Azevedo (2008; 2009) engaged their participants in 40 and 30 minute learning sessions, respectively. Although their findings had important implications for studying self-regulated learning, prior knowledge, motivation, and other educational constructs, these learning sessions may have been limited to the session time. On the contrary, van Seters et al. (2012) engaged participants in the learning session for two hours, which can be seen to some researchers as too long, and to others as an appropriate learning time. In contrast, the MetaTutor (Azevedo et al., 2012a) learning session lasts for one hour. Shapiro (2008) suggested that many different student characteristics, such as prior knowledge, can affect the way they use hypermedia environments, for example hypermedia assisted learning.

Shapiro suggests ways we can design hypermedia environments in order to promote and foster effective self-regulated learning and scaffolding while students use hypermedia assisted learning (Shapiro, 2008). It is evident, therefore, that researchers use various learning times, and so future research should continue to examine the appropriate amount of time students need in order to learn, and to engage in effective self-regulated learning processes.

All of the obtained results provide important implications into designing and testing the effect of prior knowledge on learning outcomes, and it is important for future studies to consider even more dependent variables in order to determine even more variables that can be influenced by prior knowledge in a range of learning environments. In the current study, we examine the specific processes of self-regulated learning, and how the students' prior knowledge levels can influence the use of these cognitive and metacognitive strategies.

Current Study: Overview and Hypotheses

In this study, we examined participants' pre-test scores (i.e., answers to a 25-item multiple choice test) to determine whether their prior knowledge of the human circulatory system influenced how they deployed cognitive and metacognitive SRL strategies as they interacted with MetaTutor, an intelligent, multi-agent, computer-based learning environment, where the learning goal is to learn everything one can about the human circulatory system in a 60-minute session. In addition, we used data mining techniques to examine sequential patterns of the use of SRL strategies and whether there were differences in prior knowledge groups in their use of combined sets of two product variables. We used data mining techniques to plot students' use of SRL strategies, time they spend engaging in SRL strategies, and the number of relevant pages they visited, as they engaged in various sub goals. Overall, we expected that students categorized into the high prior knowledge group would engage in more cognitive and metacognitive SRL strategies than students in the low prior knowledge group; and would, furthermore, engage in more adaptive uses of SRL strategies (e.g., a metacognitive strategy, followed by a cognitive strategy).

Methods

Participants

123[1] (66% female) undergraduate students at McGill University and Concordia University in Montreal, Quebec, and the Illinois Institute of Technology in Chicago participated in this study. Participants' ages ranged from 18 to 30 years old, with a mean age of 20.83, *SD* = 2.34. Due to some restrictions, which are based on data collection methods, participants had to meet the following criteria: (1) must have normal vision (i.e., not wear glasses or contacts, or have color vision); (2) not wear any kind of head covering, which blocks any part of the face; and (3) be able to pull hair back if it should cover the eyes or the eyebrows. Participants were given a monetary compensation of up to $40 dollars for completing the study.

Research Design

In this study, we used a quasi-experimental design, as participants were classified into either low-prior knowledge (LPK, n = 26), or high-prior knowledge (HPK, n = 26), based on their pre-test scores on the human circulatory system; with a median of test scores of 20 out of 25; M = 18.93, SD = 3.98.

[1] However, only data from 52 participants were used in this study.

Materials

Two equivalent 25-item multiple choice pretests and posttests developed by Azevedo and colleagues (Azevedo et al., 2010) were used to assess participants' learning during the one-hour learning session with MetaTutor. See Appendix A for sample questions.

We extracted data from the log files, which captured the students' interactions with the MetaTutor environment. The extracted data consisted of the SRL strategies students engaged in, as well as the time these strategies were initiated and the time engagement in the strategies ended, based on their use of the SRL palette. Additionally, we extracted sequences of SRL processes students engaged in during the learning session from the same log-files. The extracted log file data were then analyzed with respect to HPK and LPK groups.

MetaTutor: An intelligent, hypermedia multi-agent system

MetaTutor is an intelligent, hypermedia-learning environment, which engages students in learning about a complex science topic, the circulatory system (Azevedo et al., 2012, 2013). In the MetaTutor hypermedia environment, there are 38 pages of text and diagrams, all of which address different topics pertaining to the circulatory system. MetaTutor allows us to collect a wide array of data, including log-file data, eye-tracking data, think-aloud data, electro dermal activity, screen recordings of learner-system interactions, and facial expressions of participants' emotions. We collected these multi-channels of data from students while they navigated the system and learned about the circulatory system. While the participant interacted with the system, there are four pedagogical agents who are present during learning and act to assist the student in learning by providing the appropriate scaffolding for each participant. Each agent specializes in one particular area of self-regulated learning; Gavin the Guide focuses on directing the participant through the environment; Pam the Planner's role is to aid and emphasize planning, creating relevant sub-goals, and activating prior knowledge. Mary the Monitor specializes in helping the participant to keep track and monitor what is occurring throughout the learning session by emphasizing the use of monitoring progress towards goals (MPTG), content evaluation (CE), feeling of knowing (FOK), and judgment of learning (JOL). Sam the Strategizer assists in learners' use of effective strategies (e.g., creating good summaries) as they learn in the environment.

There are two conditions in this study; a *prompt and feedback condition* and a *control condition*. In the prompt and feedback condition, participants were provided with scaffolding from the pedagogical agents, and thus not required to work independently, without any assistance. In this condition, the pedagogical agents prompted the learner to engage in learning and SRL strategies, such as judgment of learning (JOL) and summarizing, in addition to learners using these strategies themselves. Furthermore, when participants in the prompt and feedback condition set their sub-goals, Pam provided feedback on their proposed sub-goals, and included feedback on whether the sub-goal was too broad or too general, and continued to assist the participants in setting the appropriate sub-goals. In the control condition, participants were free to navigate the system without any scaffolding or feedback from any of the pedagogical agents. Learners were not prompted to use any of the learning or SRL strategies that are part of the SRL palette; however they were still able to engage in and use these strategies if they chose to on their own. In addition, during the sub-goal setting phase, Pam did not provide any feedback, but simply suggested the sub-goal that the participant should choose, which he or she could decide whether or not to accept, in which case, Pam asked him or her to repeat the sub-goal, and would suggest another sub-goal for him or her to set. Therefore, in this condition, Pam simply generated a sub-goal for the participant, which required less effort and scaffolding for the participant. The participants navigated through the same environment; they were provided with the same instructions, they were presented with the same instructional videos, and they used the same multimedia learning content. The only difference between the two conditions was the scaffolding and feedback from the pedagogical agents, such that there was no assistance given from the agents to the participants in the control group. These two conditions were created, therefore, to examine the usefulness and effectiveness of pedagogical agents in scaffolding participants as they learned about a complex science topic.

The system interface includes several elements designed to detect, track, model, support, and foster self-regulated learning (see Figure 1). On the left hand side, we find the table of contents, which displays the title of each of the 38 pages. This can effectively guide the participant to select pages to read that are relevant to the current sub-goal that he or she is working on. There is a clock located above the table of contents, which informs participants of the time remaining in the session, where they will write the post-test on the circulatory system. This allows for participants to monitor the time they spend on pages, on particular sub-goals, and on learning and SRL strategies; monitoring progress towards goals (MPTG) is an important SRL strategy, which keeps learners metacognitively aware of where and how they are allotting their learning time. On the right pane of the interface, there is the SRL palette. This tool allows participants to engage in self-regulatory learning strategies, such as taking notes, judgment of learning (JOL), feeling of knowing (FOK), content evaluation (CE), creating summaries, and making inferences. Participants can use any of these SRL or learning strategies at any point during learning throughout the session, and can be either self- or agent-initiated, such that the participant can choose to engage in the learning or SRL strategies located on the SRL palette, or the agent can prompt the participant to engage in such strategies. The pedagogical agent is placed just above the SRL palette, in the top right corner of the interface. Only one agent is displayed at a time, and so depending on where the learner is in the session, learner's previous actions (e.g., metacognitive judgments regarding the relevancy of a particular page and corresponding diagram), and which learning strategies the system will provide scaffolding for, this will determine which agent is present on the screen, and this changes throughout the session. The sub-goals set by the learner are located on the top center of the interface. The sub-goal that the participant is working on is highlighted to remind the learner what he or she is working on at that moment. As the learner progresses through a sub-goal, a colored bar will display the amount of progress made, which demonstrates how far into completing the sub-goal the participant is. Furthermore, the participant can choose to complete a sub-goal at any time; the learner must then complete a quiz on the content to ensure that the participant has learned an adequate amount of information to, in fact, complete this sub-goal and move on to another one. At the bottom of the screen, there is a text box, where participants enter their sub-goals, write down their prior knowledge of the circulatory system, and of each particular sub-goal. This box is also used when the participant chooses to engage in certain SRL processes, such as judgment of learning (JOL), feeling-of-knowing (FOK), and content evaluation (CE); where he or she can select a response to questions such as, "how well do you think you understand the material presented on this page", where multiple options are given to respond with. Finally, the middle of the interface is where the text and diagrams are located, which are the materials needed in order to accomplish the overall learning goal of learning everything about the circulatory system.

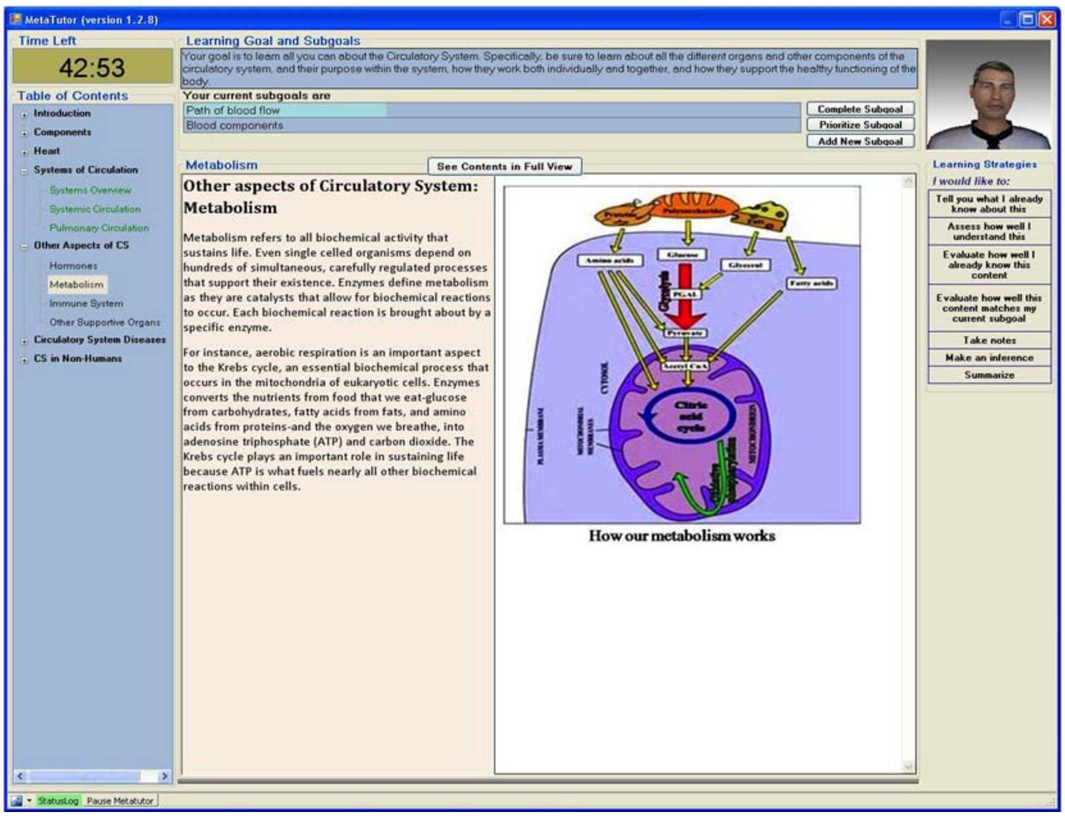

Figure 1. Screenshot of the MetaTutor Interface

Experimental Procedure

Participants had to be available for two sessions (the first for one hour and the second for three hours), which must have occurred within three days of each other. Data collection consisted of two sessions; the first session took approximately one hour, and the second session lasted up to three hours. During the first session, participants began by completing a consent form, and were then given an explanation of the study. Participants began the experiment by completing a series of self-report questionnaires, which measured demographic information and their emotions (e.g., AEQ, Pekrun et al., 2011). These emotions questionnaires were presented to the students every 14 minutes for a total of four times throughout the learning session. Participants then completed a 25-item multiple-choice pre-test to assess their prior knowledge of the human circulatory system. At the end of the session they were paid $5 for completing session one.

In the second session, participants began the session by creating sub-goals, with the assistance of Pam. In MetaTutor, there are seven pre-determined sub-goals based on different aspects of the circulatory system, which the pedagogical agents are programmed to recognize and lead the participants to set. The seven sub-goals are: (1) Path of Blood Flow, (2) Heartbeat, (3) Heart Components, (4) Blood Vessels, (5) Blood Components, (6) Purposes of the Circulatory System, and (7) Malfunctions of the Circulatory System. Once two sub-goals had been set, the participants were presented with multiple videos, which introduced the system, including all of the interface elements, and how to engage in self-regulated learning strategies, which are crucial in helping a student learn about the circulatory system (Azevedo et al., 2010, 2012). Following the introduction to the system, Pam, the pedagogical agent, asked the participant to recall all he or she knew about the circulatory system by writing it all down in the text box. Participants could write as much or as little as they knew, and so we would expect that students with a higher prior knowledge of the circulatory system would have more to present, which can be referred back to in the log-file data, which captured anything that was written into the system. Next, participants chose the sub-goal they will begin working on; and Pam asked them again to mark down everything they already knew about the given sub-goal. Finally, participants began learning with the system by freely navigating to the pages they wished to, and were able to engage in the self-regulated learning strategies at any point during the session by selecting the strategy they wished to use from the SRL palette, which they felt will be useful in helping to achieve their overall sub-goal, to learn all they can about the human circulatory system. Throughout the one-hour session, four more self-report questionnaires were presented every 14 minutes, which measured students' emotions, for a total of four times during the session, and completed by the participants before they could continue learning. The repeated administration of questionnaires were used to assess fluctuations in emotions as they progressed during the learning session with MetaTutor. When the participant completed all of his or her sub-goals, or if the one-hour learning session was up, the participant was then presented with a 25-item post-test on the human circulatory system.

During learning with MetaTutor, several multi-channel data were collected including log-files, concurrent think-aloud protocols, electro dermal activity (EDA), face expressions, eye-tracking, and audio recordings, about each participant's self-regulated behaviors. The participant was then debriefed and paid $40 for completing the study. See Appendix B for an illustration of the experimental set-up. For this thesis, I only extracted and analyzed log-file data.

Coding and Scoring: Product and Process Data

Once the participant had completed the second session, the research assistant running the session printed the log-files and placed it in the participant's file folder, re-named the video files with the corresponding participant number, plugged in and saved the EDA file, and replayed the eye-tracker to ensure that it recorded the session accurately and to its entirety. These procedures completed the experimental session and data archiving for subsequent data analyses.

In order to determine one's prior knowledge level, I conducted a median split on pretest scores, such that scores that fell below the median were categorized as low-prior knowledge (LPK) and those who scored above the median were classified as the high-prior knowledge (HPK) group. For this data set, the median score was 20 (out of 25 or 80%). In order to eliminate scores that were at the median or were too close to the median and were thus not extreme scores, scores at the median (20), scores one or two points lower than the median (18 and 19), and scores one or two points higher than the median (21 and 22) were removed from the data set, which yielded an N of 69. However, this median split yielded two groups with unequal sample sizes (n = 26 for the high prior knowledge group [HPK] and n = 43 for the low prior knowledge [LPK] group); and so, in order to create equal sized groups, I removed 17 scores from the low prior knowledge group to yield an N of 52; 26 participants per group. Therefore, scores in the high prior knowledge [HPK] group ranged from 23 to 25 (M = 23.62, SD = 0.70), and scores in the low prior knowledge [LPK] group ranged from 14 to 17 (M = 15.89, SD = 1.11). Time-stamped, log-file data were used to extract the frequency of use of SRL strategies and the time spent engaging in these SRL processes during the experimental session.

Data mining techniques were utilized in order to determine the sequences of SRL strategies, which were deployed in by participants during the session; and were compared between prior knowledge groups. More specifically, we used the following data mining techniques (Hegland, 2001; Pujari, 2001) to seek patterns in the data set: (1) anomaly detection, to identify the unusual data; (2) clustering, to structure the data; (3) regression, to determine the functions to model the data; and (4) summarization, to provide comprehensive representations of the results. The data were mapped, based on the generated hypotheses, in order to determine the influence of the SRL activities with their sequences, on HPK and LPK groups.

Furthermore, we extracted data points from three variables; number of SRL processes deployed, time spent engaging in SRL processes, and page relevancy, which were used as features in our data set. Each data set, therefore, contained three features; (1) number of SRL processes; (2) time spent on SRL processes; and (3) a binary variable for page relevancy, which were extracted from the log files, and then divided and plotted by prior knowledge group. We were then able to define the decision boundary that could formally classify these data points and enable the pedagogical agents to effectively predict the group of a new data point with respect to its prior knowledge level.

Results

In order to investigate the differences between prior knowledge groups and their use of SRL processes during learning with MetaTutor, which were based on the clicks made on the SRL palette, I conducted several analyses to test for differences between each group and each SRL strategy. The strategies were categorized as cognitive or metacognitive, which yielded two sets of multiple and parallel analyses. The cognitive strategies included in MetaTutor were taking notes (TN), making inferences (INF), creating a summary of the science content (SUMM), and prior knowledge activation (PKA). The metacognitive strategies included judgment of learning (JOL), feeling of knowing (FOK), content evaluation (CE), monitoring progress towards goals (MPTG) and monitoring (a combination of JOL, FOK, CE, PKA, and MPTG). I performed chi-squares to examine the differences in frequency distributions of these strategies among high and low prior knowledge groups.

First, I conducted a chi-square analysis to determine the difference in total SRL strategy-use between prior knowledge groups. Next, I performed several chi-square analyses to examine the differences in the total use of all cognitive strategies between prior knowledge groups, followed by a comparison between each cognitive strategy among prior knowledge groups. I then performed the same chi-square analyses for the metacognitive strategies. I began with comparing the frequencies of use of the total amount of metacognitive strategies among prior knowledge groups, followed by a comparison of the frequencies of use of each metacognitive strategy among prior knowledge groups. The next section describes all of the analyses conducted with their results and illustrative representations of the results obtained.

Research Question 1: Is there a difference in the frequency distribution of learners' use of total self-regulated learning strategies?

A chi-square analysis was performed to determine whether there was a significant difference between learners' use of the total (i.e., cognitive and metacognitive) self-regulated strategies between students with HPK or LPK. We extracted the frequencies of use of SRL strategies from the log-files, which collected data during the learning session.

A chi-square test of independence revealed that there were significant differences in the frequency distribution of learners' use of self-regulated learning strategies across prior knowledge groups; $\chi^2(1) = 10.80$, $p = .00$. See Figure 2 for an illustration of the frequency distributions of SRL strategies between prior knowledge groups; and refer to Table 1 for the numerical results obtained from this analysis. Based on these results, further analyses were performed, which differentiated between cognitive and metacognitive self-regulated learning strategies, and which are addressed in the following research questions.

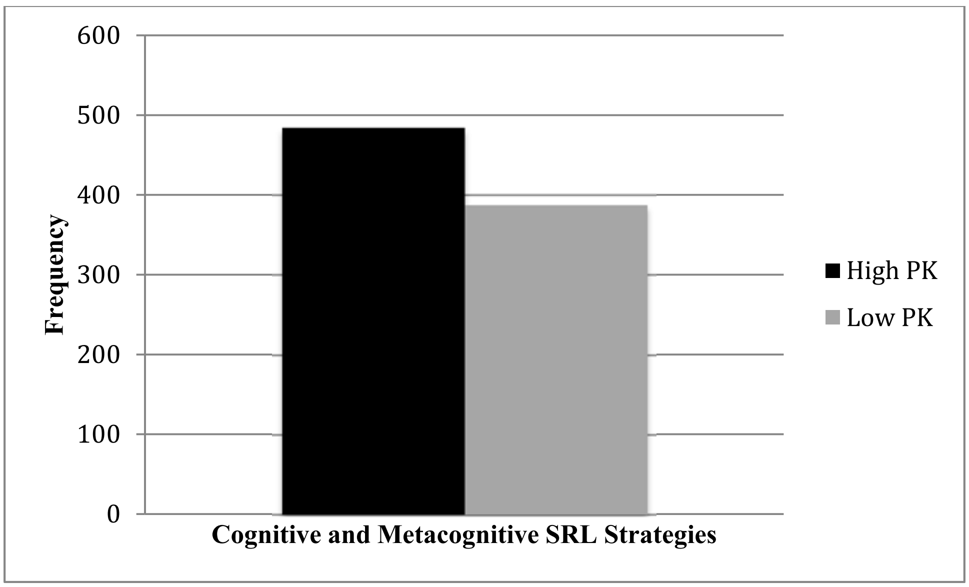

Figure 2. Frequencies of Total Self-Regulated Learning Strategies

Table 1.

Raw Frequencies and Descriptive Statistics for the use of Total SRL Strategies by High and Low Prior Knowledge Groups

	High Prior Knowledge (n = 26)			Low Prior Knowledge (n = 26)				
Variable	Raw Frequencies	*M*	*SD*	Raw Frequencies	*M*	*SD*	χ^2	*p*
Total SRL	484	2.07	4.29	387	1.65	3.16	10.80	.00*

Note: SRL = self-regulated learning.
*$p < .05$

Research Question 2: Is there a difference in the frequency distribution of learners' use of cognitive self-regulated learning strategies?

Chi-square analyses were performed to determine whether there were significant differences in the distribution of learners' use of cognitive strategies across prior knowledge groups. We extracted the frequencies from the log-files by analyzing and enumerating the uses of TN, INF, SUMM, and PKA.

A chi-square test of independence revealed that there were no significant differences in the frequency distributions of cognitive SRL strategies used across prior knowledge groups; $\chi^2(1) = 1.54$, $p = .22$. For a graphical representation of these frequencies, refer to Figure 3. A subsequent, 2X4 chi-square analysis revealed that there were no significant differences in the distributions of different cognitive SRL strategies across prior knowledge groups; $\chi^2(3) = 1.70$, $p = .06$. The frequency for each cognitive strategy, based on prior knowledge group, is illustrated in Figure 4. For an overview of the numerical results obtained from these analyses, refer to Table 2.

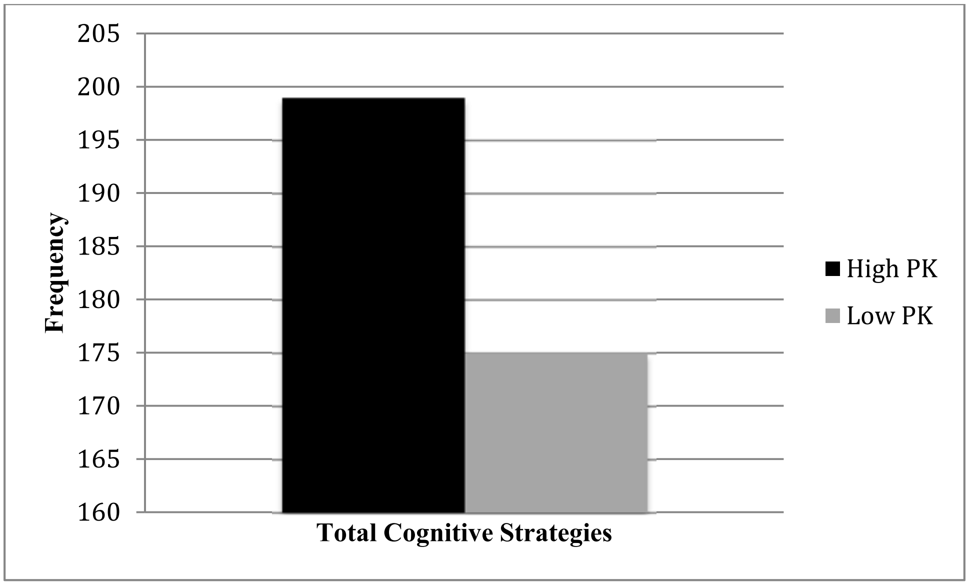

Figure 3. Frequencies of Total Cognitive Strategies

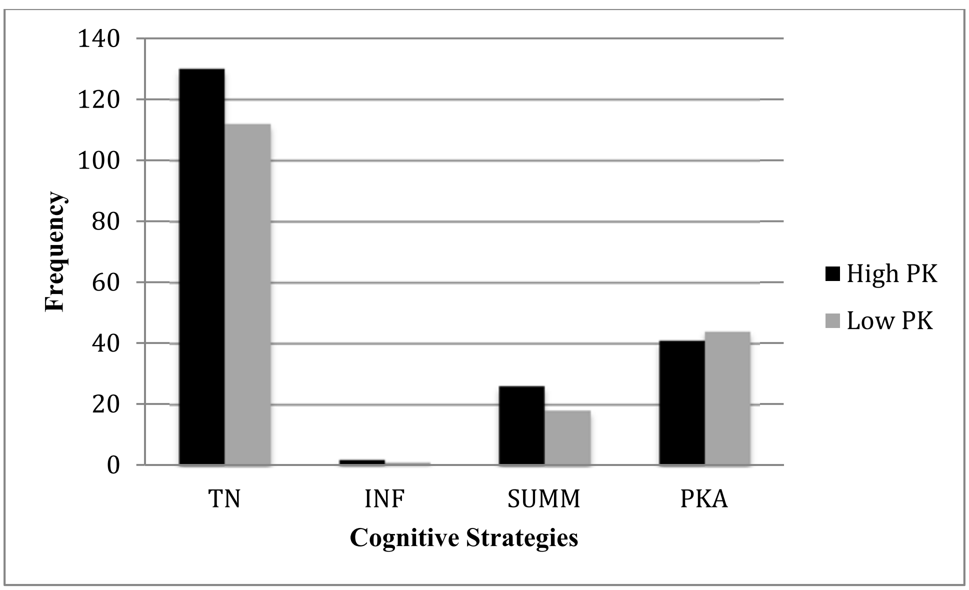

Figure 4. Frequencies of Cognitive Strategies: TN, INF, SUMM, and PKA.

Table 2.

Raw Frequencies and Descriptive Statistics for the use of Cognitive SRL Strategies by High and Low Prior Knowledge Groups

	High Prior Knowledge (n = 26)			Low Prior Knowledge (n = 26)				
Variable	Raw Frequencies	*M*	*SD*	Raw Frequencies	*M*	*SD*	χ^2	*p*
Total Cognitive	199	1.91	4.82	175	1.68	3.77	1.5	.22
TN	130	5.00	8.64	112	4.31	6.35	1.70	.06
INF	2	0.08	0.27	1	0.04	0.20	1.70	.06
SUMM	26	1.00	1.36	18	0.69	1.19	1.70	.06
PKA	41	1.58	2.16	44	1.69	2.40	1.70	.06

Note: TN = taking notes; INF = inferences; SUMM = summarizing; PKA = prior knowledge activation
*$p < .05$

Research Question 3: Is there a difference in the frequency distribution of learners' use of metacognitive self-regulated learning strategies?

Chi-square analyses were performed in order to determine whether there were significant differences in the distribution of learners' use of metacognitive strategies across prior knowledge groups. Frequencies of JOL, FOK, CE, MPTG, and Monitoring extracted from the log-files were used in these analyses.

A chi-square test for independence revealed a significant difference in the total number of metacognitive SRL strategies used across prior knowledge groups; $\chi^2(1) = 10.72$, $p = .00$. Results indicated that HPK learners used significantly more metacognitive strategies that LPK learners. The frequencies of the total metacognitive strategies used among both prior knowledge groups are displayed in Figure 5. A 2X5 chi-square analysis revealed no significant differences between the distribution of metacognitive SRL strategies used across prior knowledge groups; $\chi^2(4) = 6.26$, $p = .18$. See Figure 6 for the frequencies of each of the metacognitive strategies by prior knowledge group. Refer to Table 3 for all numerical results obtained from these analyses.

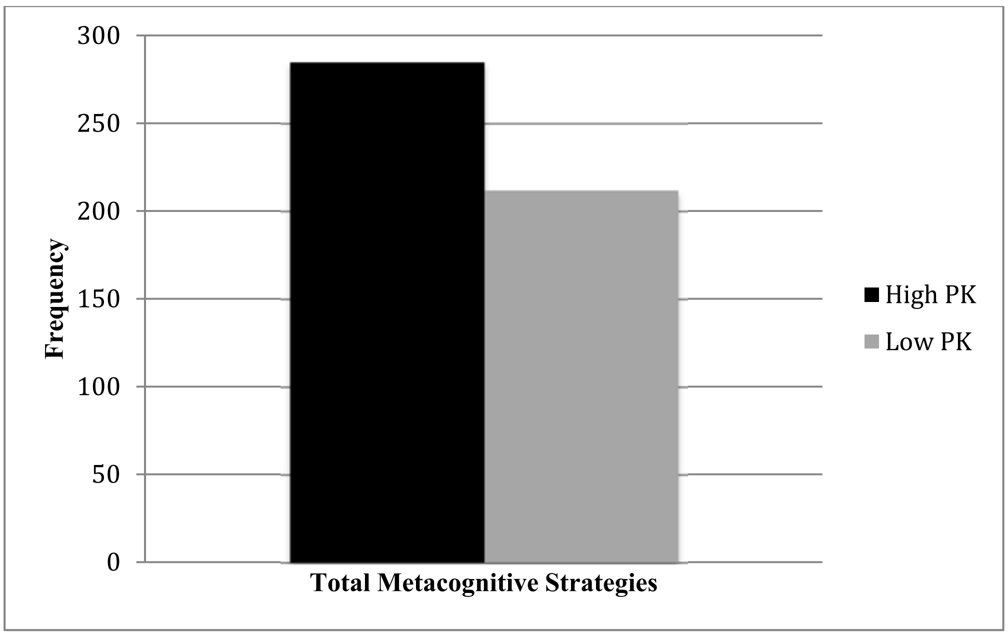

Figure 5: Frequencies of Total Metacognitive Strategies.

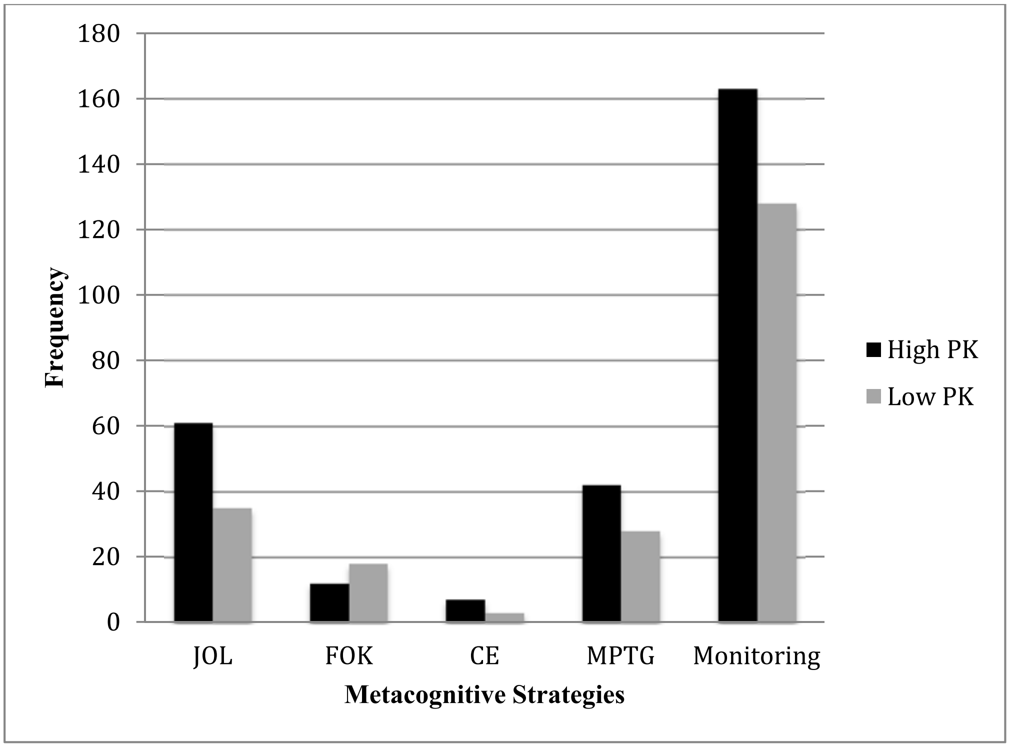

Figure 6. Frequencies of Metacognitive Strategies: JOL, FOK, CE, MPTG and Monitoring.

Table 3.

Raw Frequencies and Descriptive Statistics for the use of Metacognitive SRL Strategies by High and Low Prior Knowledge Groups

	High Prior Knowledge (n = 26)			Low Prior Knowledge (n = 26)				
Variable	Raw Frequencies	*M*	*SD*	Raw Frequencies	*M*	*SD*	χ^2	*p*
Total Metacognitive	285	2.19	3.84	212	1.63	2.58	10.72	0.00*
JOL	61	2.35	4.10	35	1.35	1.94	6.26	0.18
FOK	12	0.46	1,10	18	0.69	1.19	6.26	0.18
CE	7	0.27	0.67	3	0.12	0.43	6.26	0.18
MPTG	42	1.62	1.10	28	1.08	1.44	6.26	0.18
Monitoring	163	6.27	5.63	128	4.92	3.44	6.26	0.18

Note: JOL = judgment of learning; FOK = feeling of knowing; CE = content evaluation; MPTG = monitoring progress towards goals.
*p < .05

Research Question 4: What are the most frequent sequences of SRL strategies that differentiate between prior knowledge?

In order to answer this research question, we used the data mining techniques, including anomaly detection for the identification of unusual data, clustering for structuring the data, regression for finding functions that can model the data, and summarization for providing the compact representation of the results to seek patterns in the data set (Berkhin, 2006; Hegland, 2001). We extracted the most commonly used quintet sequences of cognitive and metacognitive SRL strategies throughout the learning session based on levels of prior knowledge. We extracted a set of five sequences because they represent an adequate time-scale of learner-system interactions from which to interpret SRL behaviors during learning with MetaTutor. In an ITS, the goal for the agent is to scaffold the student to learn in the most efficient way, in order to accomplish sub goals, and engage in effective cognitive and metacognitive processes. The agent does this by assessing the student's performance as he or she interacts with the system. If the agents are programmed to assess how the students engage in sequences of SRL processes, it might not be beneficial to program agents to assess a large sequence, because the student might have benefitted from agent scaffolds and feedback sooner in the session. Therefore, we used five sequences of SRL strategies to better the agents' planning capabilities, so they can investigate the most effective patterns to guide the students of different prior knowledge levels. Quintet sequences were categorized into seats, such that a seat represented the position within the sequence (e.g., seat 1 is position 1 of the sequence of 5). Overall, results indicated that for the HPK group, the most common seats in a quintet sequence, based on descending percentage of frequency use was: (1) prior knowledge activation (PKA) in seat one and used by 10.11% of HPK learners; (2) taking notes (TN) was engaged in seat two and used by 11.07% of HPK learners; (3) judgment of learning (JOL) was engaged in seat 3 by 9.30% of participants with HPK; (4) feeling of knowing (FOK) was engaged in seat 4, by 9.04% of HPK participants; and (5) monitoring progress towards goals (MPTG) was engaged in seat 5 by 8.34% of participants in the HPK group. For a summary of the percentages of the four most frequent SRL strategies used in the quintet sequences by the participants in the HPK group, refer to Figure 7.

In contrast, the quintet sequences used by LPK participants illustrates the use of different sequences of SRL strategies. While 9.22% of participants with LPK did engage in PKA in seat 1, unlike participants with HPK, they did not engage in the same SRL strategies most frequently in seats two through five. 10.695% of participants with LPK engaged in monitoring progress towards goals (MPTG) in seat two, followed by 10.52% of LPK students, who engaged in feeling of knowing (FOK) in seat three, 10.10% of participants who engaged in JOL in seat four, and 8.57% engaged in MPTG in seat five. For a graphical representation of the percentages of the four most frequent SRL strategies of the quintet sequences students with low prior knowledge engaged in, see Figure 7.

Based on these obtained results, PKA is the most frequently used SRL strategy in seat one of five of the quintet sequence across both groups. In comparing seats two through five, however, students with HPK most frequently engaged in cognitive SRL strategies (e.g., taking notes) in seat 2, while students with LPK engaged in cognitive strategies (e.g., SUMM) most frequently in seat 5. HPK students engaged in metacognitive SRL strategies (e.g., JOL, FOK, and MPTG) most frequently in seats 3 through 5, while students with LPK most frequently engaged in metacognitive SRL strategies (e.g., MPTG, FOK, and JOL) in seats 2 through 4. To further investigate the quintet sequences used by participants, an analysis, which involved clustering of the data, followed by the regression process to formalize the patterns that were found of the most frequently used SRL strategy quintet sequences and least frequently used quintet sequences of SRL strategies, by HPK and LPK students during the learning session, were assessed and are presented in the next paragraph.

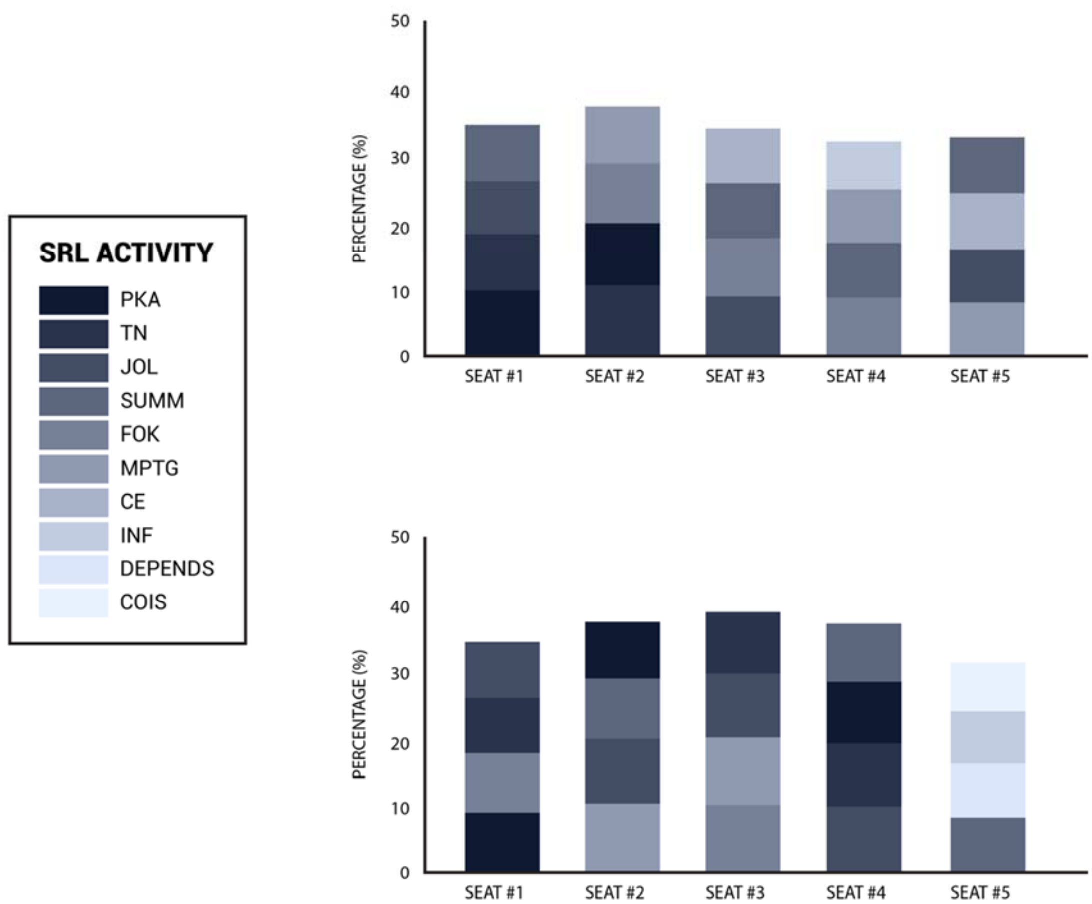

Figure 7. Percentages of the four most frequent SRL strategies for HPK (top) and LPK (bottom) groups

Note: PKA = prior knowledge activation; TN = taking notes; JOL = judgment of learning; SUMM = summarizing; FOK = feeling of knowing; MPTG = monitoring progress towards goals; CE = content evaluation; INF = making inferences; DEPENDS = when a student leaves a page after more than 5 seconds, but less than 14 seconds, and the agent prompts the student to explain why this is so; and COIS = coordination of informational sources.

Most frequently deployed SRL sequences. I used data mining techniques to examine over 100,000 quintet sequences of SRL strategies collected during the entire learning session from each of the 52 HPK and LPK participants. Results demonstrated that some sequences were engaged in 44 times, while others were engaged in one or zero times. For example, HPK learners engaged in the sequence of: PLAN-PKA-JOL-MPTG-SUMM 44 times, which was the most frequent sequence across the HPK sample. Next, HPK learners engaged in PLAN-PKA-JOL-SUMM-TN 43 times, followed by PLAN-PKA-JOL-MPTG-CE and PLAN-PKA-SUMM-TN-MPTG, which had both been engaged in 42 times. It is evident, therefore, that students with HPK most frequently engaged in PLAN, then PKA, as the first two SRL strategies. In the four most frequent quintet sequences of SRL strategies used by HPK students, they began with these two SRL strategies. The most frequent sequence; PLAN-PKA-JOL-MPTG-SUMM, involved students creating summaries after monitoring their progress towards goals, which is an effective sequence of strategies, since they decided to create summaries after they planned, activated their prior knowledge of the content, judged whether or not they understood what they were learning, and monitored how far along they were to completing their sub-goals. In the second most common sequence; PLAN-PKA-JOL-SUMM-TN, learners created summaries and took notes after they planned, activated their prior knowledge, and judged whether or not the content was relevant, which is also an effective sequence, since they metacognitively planned and monitored their progress, and then engaged in the appropriate cognitive learning strategies. The students in the LPK group differed in the sequences of SRL strategies they used, however they also used some similar sequences of SRL strategies. Like those in the HPK group, the four most frequent quintet sequences of SRL strategies used by LPK students began with planning, followed by PKA. The two most common sequences used were PLAN-PKA-SUMM-TN-FOK and PLAN-PKA-SUMM-TN-JOL, which were both used 42 times. The next two most frequent quintet sequences used by LPK students were PLAN-PKA-JOL-SUMM-

TN and PLAN-PKA-SUMM-TN-MPTG, which were both used 41 times. These sequences were the same as those used by HPK participants, which demonstrates that although we do see differences in the sequences of use of SRL strategies between prior knowledge groups, participants with LPK are not all necessarily ineffective in using SRL strategies. Previous results indicated no significant differences between the frequencies of use of some SRL strategies, and so it is not surprising to find results, which demonstrate that students in both prior knowledge groups engage in similar sequences of use of SRL strategies.

Least frequently deployed SRL sequences. In addition to the most frequent quintet sequences of SRL strategies used, we also identified sequences that were rarely used by students during the learning session. Many of the sequences were not used at all, however we will assess some of the sequences that were only used once in each prior knowledge group. In the HPK group, there were many quintet sequences, which were used once, and so we will only focus on some of them. One sequence, INF-RR-MPTG-JOL-FOK, may not have been used by participants because it is often beneficial to judge one’s learning (JOL) and assess if this material had been previously seen (FOK) prior to re-reading the content or making inferences. In order to make inferences, it is important for students to ensure they understand the content; and if they do not understand, they should re-read the page. It would be ideal for both of these processes to occur prior to making inferences about the content. A large number of sequences that began with making inferences, followed by re-reading were only used once in the HPK group, which makes sense because it is more beneficial for students to make inferences and re-read content after they engage in important metacognitive strategies, which can be used to plan and monitor further use of SRL strategies during the learning session. Another sequence, which one student in the HPK group used, was RR-PLAN-TN-FOK-SUMM, which is not a sequence that we would expect to see from students with high prior knowledge. Students with high prior knowledge should not need to re-read the content because they are already familiar with it. It is, furthermore, not ideal for students to determine whether or not they have seen the content after they have taken notes on it. It may be beneficial for them to summarize the content after they had determined they had seen the content before (FOK), however the strategies that precede the FOK and SUMM are not ideal for high prior knowledge students to engage in. We would expect that students would engage in an FOK closer to the beginning of a sequence of SRL strategies, because they will need to assess if the content is familiar before they select the appropriate SRL strategies to use.

Students with low prior knowledge, similar to students in the HPK group, rarely engaged in many sequences of SRL strategies. One sequence, which was only engaged in once during the learning session, was INF-RR-COIS-CE-JOL. This seems appropriate because it is not beneficial to make an inference prior to re-reading content (RR) and coordinating the content with the images (COIS). In order for an inference to be made, students should grasp an understanding of the content. Furthermore, the evaluation of the content's relevancy to the sub goal (CE) should occur before making an inference, re-reading, or coordinating informational sources, for if the content is not relevant to the current sub goal, it is not beneficial to make an inference on the content, nor should he or she read the content over again, and proceed to coordinate the diagrams with the text. In addition, a student should judge whether or not he or she understands the content before engaging in cognitive processes, such as making inferences and re-reading. Another sequence, which was only used once during learning, was CE-INF-FOK-SUMM-MPTG. This can be seen as an effective SRL quintet sequence because it begins with a student evaluating the relevancy of the content, which is important to consider when navigating to different pages, some of which are not relevant to the current sub goal, followed by making an inference, which is a complex cognitive learning strategy. These two strategies are followed by FOK, SUMM, and MPTG, which are all effective SRL strategies in the order they were used in. It can be helpful to assess whether or not the content has been seen before, and we would expect that if someone has seen this content, it can be followed by writing down the information, for example, creating a summary of the content, which we expect that students with LPK would not do, since they might want to read the content first. Finally, it is also beneficial to monitor one's progress towards completing a sub goal after engaging in cognitive strategies, and so engaging in MPTG is an effective strategy to engage in at the end of a sequence of strategies. We may not expect students with LPK to engage in this sequence because it implies effective use of SRL. It has previously been mentioned that students with LPK may not have the capacity to engage in effective SRL because they are allotting their time to understanding content.

Overall, the data mining sequences has shown that students with HPK engage in different sequences of SRL strategies than students with LPK; however, we also did see some similarities in the sequences used by both prior knowledge groups, which seems appropriate since our previous results demonstrated that there were some significant and non-significant differences in the frequencies of use of SRL strategies among prior knowledge groups. There are many sequences of strategies, which have not been assessed in this thesis, however it is important to remember that there are many SRL sequences, which should be examined and analyzed in a similar manner.

Research Question 5: Are there differences between learners' time and use of SRL strategies across individual and combined MetaTutor learning sub-goals, based on knowledge groups?

In order to address this research question, we used the data mining techniques (Berkhin, 2006; Hegland, 2001; Pujari, 2001); anomaly detection, clustering, regression, and summarization, which combined sets of product variables, such as time engaging in SRL processes, in order to plot individual data sets, which were categorized by sub-goal. The values for these variables were obtained by extracting the information from the log-files. In MetaTutor, an SRL activity is considered as an event. When a student clicks on the SRL palette, the associated event is activated and therefore, the initiation time is recorded. Likewise, the deactivation time is also recorded in the log-file. There might be no deactivation for an SRL process, but the initiation of a new SRL process would automatically deactivate the previous one. The variables used for analyses in this research question were the number of SRL processes engaged in and time spent engaging in SRL processes. HPK students are shown as blue dots, and LPK students are shown as red dots. Figure 8 displays the plots of different prior knowledge students' data points with respect to the number of SRL processes used and the time spent engaging in SRL processes for each of the seven MetaTutor sub-goals (see list on page 19 of the thesis), and for all of the seven sub-goals combined, for a total of eight data plots.

For the *Path of Blood Flow* MetaTutor sub-goal, HPK students engaged in more SRL processes and spent more time engaging in SRL processes, compared to students with LPK. For the *Heartbeat* MetaTutor sub-goal, in general, students in both groups did not engage in as many SRL processes as seen used in other sub goals, and therefore did not spend much time engaging in SRL processes. For *Heart Components*, students in the HPK group spent more time engaging in SRL processes and they engaged in more SRL processes. According to Figure 8, few students in the LPK group selected *Heart Components* as the sub-goal, or they did not engage in any SRL processes while completing this sub-goal. For the *Blood Vessels* MetaTutor sub-goal, students in the LPK group engaged in more SRL processes, where the students who spent more time engaging in SRL processes engaged in a larger number of SRL processes than students who spent less time engaging in SRL processes. For *Blood Components*, students with HPK engaged in more SRL strategies, and thus spent more time engaging in SRL strategies than students with low prior knowledge, which implies that there seems to be more students with HPK who set blood components as a sub-goal, and engaged in SRL strategies, compared to students with LPK, who rarely set *Blood Components* as a sub-goal, and those who did used less SRL strategies than students with high prior knowledge. We see a similar pattern with *Purposes of the Circulatory System*, where there are more HPK students, and those that engaged in more SRL strategies spent more time engaging in SRL processes throughout the session. Many students with LPK did not engage in many SRL processes and thus did not spend much time engaging in SRL processes, with the exception of a select few, who overall did spend more time engaging in SRL processes if they engaged in more SRL processes. Finally, for *Malfunctions of the Circulatory System*, there were more HPK students who set this sub-goal during the second session of the experiment than students with low prior knowledge. Furthermore, the students with HPK who spent more time engaging in SRL processes were the ones who engaged in more SRL processes for the sub-goal. The few students with LPK who used SRL strategies used a small number of SRL processes, and thus did not spend much time engaging in SRL processes. From these results, we can infer that students with LPK either did not engage in SRL processes as they attempted to complete this sub-goal, or they did not set *Malfunctions of the Circulatory System* as one of their sub-goals in the learning session. Overall, the results from this analysis revealed that there appears to be differences in the sub-goals that were set by HPK students, compared to LPK students, and that we see different uses of SRL strategies, and durations of uses in engaging in SRL processes depending on the sub-goal that is being attended to *Heart Components*, *Purposes of the Circulatory System*, and *Malfunctions of the Circulatory System* seem to be the sub-goals where students with HPK engaged in a larger number of SRL processes compared to the number of SRL processes they used when working on the other four sub goals; *Path of Blood Flow*, *Heartbeat*, *Blood Vessels*, and *Blood Components*. There seems to be, overall, a lower number of SRL processes used by LPK students, which implies that these students are engaging in less SRL processes during the learning session than students with HPK. Refer to Appendix C for Tables, which display the X and Y-axes coordinates in numerical values of each student's data plot.

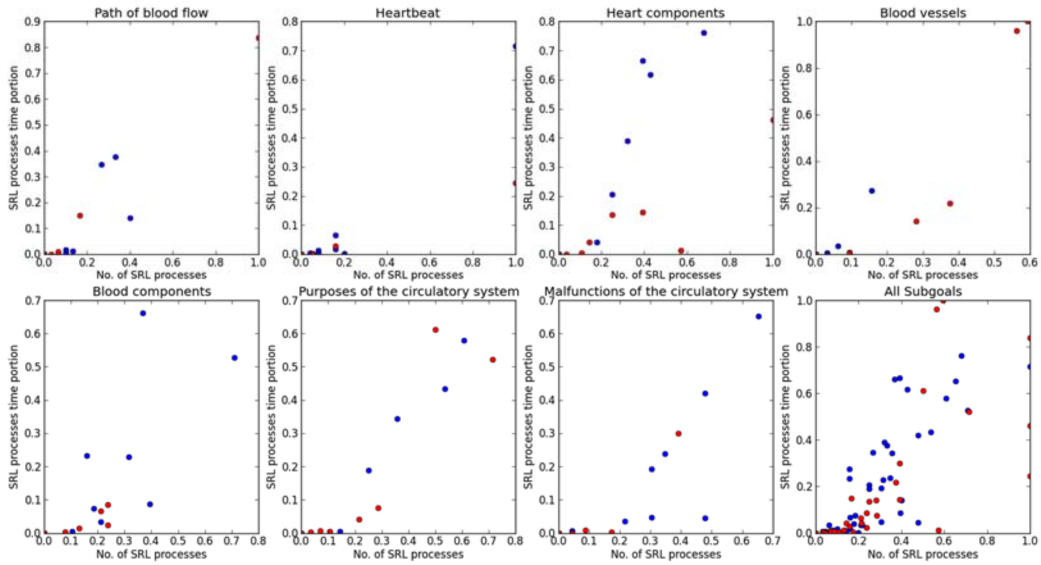

Figure 8. Data Plots for each of the Seven MetaTutor Sub-Goals and All Seven Sub-Goals based on the Number of SRL Processes and Time Spent Engaging in SRL Processes

Research Question 6: Are there differences between learners' use of SRL strategies and visits to relevant sub goal pages across individual and combined MetaTutor learning sub-goals, based on knowledge groups?

Another pair of product variables, which were assessed by individual's data plots, were the number of SRL processes with page relevancy. Refer to Research Question #5 for the details in extracting the number of SRL and time spent on SRL variables from the log files. In order to extract the page relevancy data, the log-file records the pages that the participant read, as well as the sub goals that he or she set during learning. Moreover, we predetermined which pages were relevant to each sub goal; and so, we calculated the ratio using the current sub goal and the relevant pages read throughout accomplishing the sub goal. See Figure 9 for student plots of SRL processes with page relevancy, for each sub goal and for all sub goals. Similar to Figure 8, each plot represents a data point for each participant; and the HPK students are represented in blue, while the LPK students are represented in red. For the *Path of Blood Flow* sub-goal, students with HPK and LPK both appeared to have engaged in this sub-goal and engaged in SRL processes, however there does appear to be a higher use of SRL processes with higher page relevancy among students with HPK, compared to students with LPK, who seemed to have engaged in fewer SRL processes. For the *Heartbeat* sub-goal, students with both HPK and LPK engaged in few SRL processes, however students with HPK visited more relevant pages than students with LPK. For *Heart Components*, overall, HPK students engaged in more SRL processes than LPK students, however all students visited the same number of relevant pages, regardless of the amount of SRL processes engaged in. Students with both HPK and LPK, who visited in the most relevant pages did not, however, engage in the most SRL processes, which can be inferred as students who are engaging in less, but more effective SRL processes, are visiting the most relevant pages. For the *Blood Vessels* sub-goal, students with HPK engaged in fewer SRL processes and visited slightly fewer relevant pages than LPK students. For *Blood Components*, students with HPK engaged in more SRL processes than students with LPK; however, one student with LPK visited more relevant pages than the HPK cluster of students, and one student with LPK visited the same amount of relevant pages as students with HPK, which further demonstrates the abovementioned results of a lack of significant differences in the use of SRL strategies among prior knowledge groups. It can also be noted that there were many more students with high prior knowledge who engaged in this sub-goal, or who engaged in SRL processes for this sub goal. For *Purposes of the Circulatory System* sub-goal, results indicated that more LPK students engaged in this sub-goal, which can be explained by fewer students with HPK engaging in this sub-goal, or these HPK students did not engage in SRL processes for this sub goal. Moreover, LPK students engaged in a range of number of SRL processes, and visited similar numbers of relevant pages, compared to students with HPK, who, generally, visited more relevant pages and engaged in more SRL processes than LPK students. Finally, for the *Malfunctions of the Circulatory System* sub-goal, there were far more HPK students who engaged in this sub-goal. Furthermore, these HPK students engaged in more SRL processes, and visited fewer relevant pages than LPK students. It is, however, important to consider that more LPK students may not have engaged in SRL processes for this sub-goal, which is why there appears to be far less LPK students who engaged in this sub-goal. This sub-goal can be seen as more difficult than other sub-goals, and so it is probable that LPK students did not choose to engage in this sub-goal because they were less familiar in setting it, or they may have been less familiar with the content or knowledge related to this challenging sub goal.

Overall, students appeared to engage in a varying number of SRL processes, while visiting relatively similar numbers of relevant pages. More specifically, regardless of what sub-goal the students were working on, they seemed to navigate to the same number of pages, which were relevant to their sub goal, during learning. However, students tended to use differing numbers of SRL processes depending on the sub-goal being worked on. Students with both HPK and LPK used lower numbers of SRL processes for *Path of Blood Flow*, and *Heartbeat*; and most students used more SRL processes for *Heart Components*, *Blood Vessels, Blood Components, Purposes of the Circulatory System,* and *Malfunctions of the Circulatory System*. *Path of Blood Flow* and *Heartbeat* can be seen as easier sub-goals, and so we might assume that fewer SRL processes would be needed in order to complete these sub-goals. Refer to Appendix D for

Tables, which display the numerical values of X and Y-axes coordinates of each student's data plot.

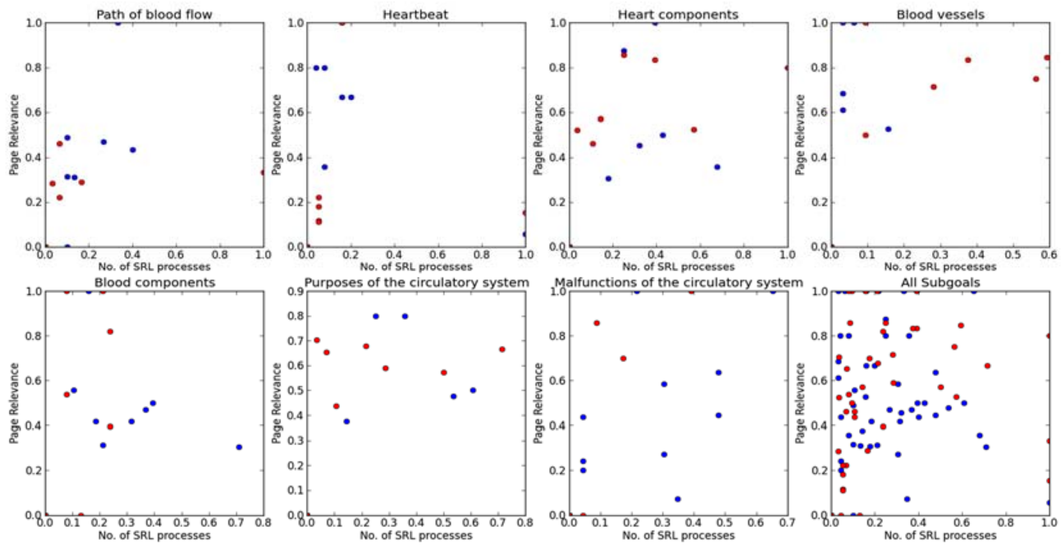

Figure 9. Data Plots for each of the Seven MetaTutor Sub Goal and All Seven Sub Goals based on the Number of SRL Processes and Page Relevancy

Discussion

The results from this study demonstrated how low- and high-prior knowledge students used cognitive and metacognitive SRL strategies as they learned about the human circulatory system in a 60-minute session with MetaTutor, a multi-agent intelligent hypermedia system. More specifically, we investigated how students' prior knowledge of the circulatory system affected how they used different learning strategies, such as taking notes, prior knowledge activation, judgment of learning, feeling of knowing, and others. Results indicated that prior knowledge groups significantly differed in their total use of SRL strategies, and total use of metacognitive SRL strategies; however prior knowledge groups did not significantly differ in their frequencies of use of total cognitive SRL strategies, nor did they differ in the use of individual metacognitive or cognitive SRL strategies. Furthermore, prior knowledge groups appeared to differ in their sequences of use of SRL strategies and their engagement in SRL strategies as they engaged in the MetaTutor sub-goals. Students did not appear to differ in the amount of pages they visited, which were relevant to the sub goals they were working on. The following sections will address the specific results obtained, based on each research question.

Research Question 1: Is there a difference in the frequency distribution of learners' use of total self-regulated learning strategies?

This question addressed the overall frequencies of use of self-regulated learning strategies, such that it determined if students with different levels of prior knowledge differed in their use of all self-regulated learning strategies as they learned with the MetaTutor environment. Results demonstrated that the prior knowledge groups did differ significantly in the total use of SRL strategies, which supports the first hypothesis that suggested that students with different prior knowledge would differ significantly in their use of self-regulated learning strategies. Furthermore, these results support the majority of findings in prior knowledge research, such as those by Shapiro (2004), Winters and Azevedo (2005), Scheiter et al (2009), and Moos and Azevedo (2008; 2009), who all found a significant effect of prior knowledge on student learning. Thus, these findings further emphasize the importance of prior knowledge in learning, and how we should consider students' prior knowledge levels when designing hypermedia-learning environments.

Research Question 2: Is there a difference in the frequency distribution of learners' use of cognitive self-regulated learning strategies?

This question addressed the use of cognitive SRL strategies among prior knowledge groups, by using chi-square analyses, to determine if there were significant differences between HPK and LPK groups in their frequency distributions of taking notes, making inference, creating summaries, and activating prior knowledge; and of the total use of these strategies. Results indicated no significant differences, such that prior knowledge groups did not differ based on their use of total cognitive strategies, nor did they differ in their use of the specific cognitive strategies (e.g., TN, INF, SUMM) and PKA, which can be used during the MetaTutor learning session. These results, therefore, do not support the hypothesis that HPK students would engage in more cognitive SRL strategies than LPK students; and which thus support Shapiro's (1999) and van Seters et al.'s (2012) findings, where they found that prior knowledge did not influence students' learning goals and did not affect students' learning paths. These obtained findings, therefore, contribute to SRL research, such that when creating environments that adapt to prior knowledge levels, researchers may want to focus on aspects related to SRL other than cognitive strategy use.

Research Question 3: Is there a difference in the frequency distribution of learners' use of metacognitive self-regulated learning strategies?

This research question addressed if there were significant differences in the total use of metacognitive strategies and differences in the use of the individual metacognitive strategies JOL, FOK, CE, MPTG, and Monitoring, among prior knowledge groups, as they engaged in learning with MetaTutor. The results demonstrated that although there were no significant differences in the use of each metacognitive strategy, there were significant differences in the total use of metacognitive SRL strategies between HPK and LPK students. This partially supports the initial hypotheses, which stated that HPK students would engage in more metacognitive processes than LPK students, since we did see a significant difference among prior knowledge groups; however this difference was only found in the total use of metacognitive strategies, and no differences were found between the specific metacognitive strategies. Thus, the results, which reported a significant difference in the frequency distribution of the total metacognitive strategies between prior knowledge groups support findings reported by Shapiro (2004), Winters and Azevedo (2005), Scheiter et al (2009), and Moos and Azevedo (2008; 2009), who found a significant effect of prior knowledge on student learning. Results, which did not find a significant difference between individual metacognitive strategies between prior knowledge groups support findings reported by Shapiro (1999) and van Seters et al (2012), who did not find a significant effect of prior knowledge on learning outcomes. These findings can contribute to research in SRL because they emphasize the level of granularity at which SRL processes are coded and analyzed. More specifically, it is easier to obtain a significant difference when all processes are clumped at the macro-level and therefore becomes more difficult when analyses are conducted at a micro-level (e.g., JOL, FOK) because the raw frequencies become diluted and thus lead to no statistically significant results. However, from a theoretical and design perspective, it is important to analyze the data at a micro-level.

Research Question 4: What are the most frequent sequences of SRL strategies that differentiate between prior knowledge groups?

This research question assessed what were the most commonly used quintet sequences of SRL strategies, and whether HPK and LPK students differed in their most frequent use of these SRL strategies. Results indicated that all students used prior knowledge activation (PKA) the most frequently in seat one of five; however students with HPK engaged in TN, JOL, FOK, and MPTG most frequently in seats two through five, respectively; and students with LPK engaged in MPTG, FOK, JOL, and SUMM in seats two through five, respectively. This suggests, therefore, that HPK students engaged in a cognitive SRL process the most frequently in seat two, followed by metacognitive SRL processes most frequently in seats three through five, while LPK students engaged in metacognitive processes most frequently in seats two through four, followed by most frequently engaging in a cognitive process in seat five. It can be presumed that students with HPK engaged in more cognitive processes because they did not physically engage in the metacognitive processes on the SRL palette; and so this data was not captured by the log-file data. HPK students may have engaged in accurate metacognitive judgments so rapidly that they were covert and did not necessitate the clicking of metacognitive processes such as JOL and FOK on the SRL palette. Evidence of this was found in results obtained by Moos & Azevedo (2008, 2009).

Furthermore, the most frequent sequence for HPK students was PLAN-PKA-JOL-SUMM-TN, while LPK students most frequently engaged in PLAN-PKA-SUMM-TN-FOK. These results, thus, demonstrate that HPK and LPK students most often engaged in different sequences of SRL processes. HPK students engaged in metacognitive strategies before cognitive strategies because they were more focused on monitoring what they knew from what they did not know, and this requires metacognitive knowledge and skills; and they also had more working memory capacity to allocate to metacognitive monitoring processes. LPK students engaged in cognitive strategies before metacognitive strategies because they were focused on learning the material, therefore using more cognitive strategies, which supports the hypothesis that HPK students would be able to engage in more effective uses of SRL processes (i.e., metacognitive prior to cognitive) than LPK students. According to previously mentioned analyses, there were some significant differences among prior knowledge groups in the frequency of use of metacognitive SRL processes, and so these results were expected. We did, however, discover sequences, which were similar among both knowledge groups, and this can be further explained by the lack of significant differences in the uses of cognitive strategies among prior knowledge groups. These results, therefore, support the findings made by Shapiro (2004), Winters and Azevedo (2005), Scheiter et al (2009), and Moos and Azevedo (2008; 2009), who found a significant effect of prior knowledge on learning; however the results obtained by Shapiro (1999) and van Seters et al (2012) were also supported, since we sometimes did not find differences in learning among prior knowledge groups. These results can, therefore, contribute to research in this field because it encourages researchers to seek where students with differing prior knowledge levels differ when interacting with hypermedia-learning environments, in order to design the most effective CBLEs to promote effective learning in students with all levels of prior knowledge.

Research Question 5: Are there differences between learners' time and use of SRL strategies across individual and combined MetaTutor learning sub-goals based on knowledge groups?

This research question sought to determine if there would be differences in the duration of and use of SRL strategies between prior knowledge groups, as they worked on the MetaTutor sub-goals. Overall results indicated that duration of and use of SRL processes seemed to be lower for LPK students compared to HPK students; however the results also depended on the sub goal that was being worked on. Students with HPK appeared to set *Heart Components, Purposes of the Circulatory System,* and *Malfunctions of the Circulatory System* more frequently as their sub goals, compared to the other sub-goals. It can be inferred, therefore, that few students with high prior knowledge set *Path of Blood Flow, Heartbeat, Blood Vessels,* or *Blood Components* as a sub-goal, and those that did, did not engage in many SRL processes, and thus did not spend time engaging in SRL processes. These results support the hypothesis, which stated that HPK students would engage in and spend more time engaging in more SRL processes than LPK students. Such results, therefore, support findings reported by Shapiro (2004), Winters and Azevedo (2005), Scheiter et al (2009), and Moos and Azevedo (2008; 2009), who found a significant effect of prior knowledge on learning. It should be noted, however, that HPK students and LPK students, at times, engaged in different sub goals, and so these results can contribute to research in this field by influencing the way we program multi-agent systems to adapt to students' prior knowledge levels by providing different sub goals for them to work on.

Research Question 6: Are there differences between learners' use of SRL strategies and visits to relevant sub goal pages across individual and combined MetaTutor learning sub-goals based on knowledge groups?

This research question examined if there were differences between prior knowledge groups based on the use of SRL strategies and navigation to relevant pages, as students worked on MetaTutor sub goals. Results indicated that regardless of the sub goal students were working on, all students visited the same relevancy of pages. We did see differences, however, in the number of SRL processes students used for the different sub goals. Both HPK and LPK students engaged in fewer SRL strategies for *Path of Blood Flow* and *Heartbeat*, and engaged in more SRL processes for the other sub goals. We can assume that these are easier sub goals, and thus require students to use fewer SRL strategies as they worked on these sub goals. We did not see many differences between knowledge groups, which seems appropriate based on previously mentioned findings that did not find significant differences in the frequency distributions of cognitive strategies between prior knowledge groups. For example, for the *Blood Vessels* sub-goal, students with HPK engaged in fewer SRL processes and visited slightly fewer relevant pages than LPK students, which is not a result that we would expect. However, due to previously mentioned results, there were few significant differences in the use of SRL processes between students with high and low prior knowledge, which is how these results can be explained. The results from this research question, therefore, support findings made by Shapiro (2004), Winters and Azevedo (2005), Scheiter et al (2009), and Moos and Azevedo (2008; 2009), who found a significant effect of prior knowledge on learning; and also support findings made by Shapiro (1999) and van Seters et al (2012), who did not find a significant effect of prior knowledge on learning outcomes. Therefore, these results contribute to research in the field by emphasizing the importance of setting sub-goals when working in a hypermedia-learning environment, and that we need to adapt learning environments, which contain the appropriate sub-goals for students to learn effectively, based on their individual learning needs.

Limitations of the Study

There were several limitations to this study, which may have influenced the data that were collected and the results that were obtained. We determined prior knowledge by assessing pretest scores, which were measured by students' performance on a multiple-choice test. Thus, students who did not know the response could have gotten it correct by chance, and so may have scored high on the test, but could have low prior knowledge on the content. Moreover, this study used a subset of participants from a larger study based on a median-split and therefore the results are sample-specific to the subset included in this thesis. In addition, the participants in the study were in both experimental conditions (i.e., prompt and feedback, and control), which differ based on the prompts given by the pedagogical agents. Participants in the prompt and feedback condition received prompts from the agents to engage in a number of SRL strategies, while the agents did not prompt students in the control condition. Therefore, students in the prompt and feedback condition, regardless of their prior knowledge, may have used higher frequencies of SRL strategies throughout the learning session because they were instructed to do so. Finally, this study only included data obtained from participants' log-files, which limited the data mining analyses, which generated qualitative results. The log-files also limited our findings because participants may have engaged in other processes that were covert, and thus were not captured from the log-files. For example, HPK students may have engaged in more PKAs and FOKs, but they may have done so out loud, and did not select the option to do so on the SRL palette. The log file data, therefore, would not have captured the use of these strategies, and so we would have had to watch the video recordings to observe this. It is important, therefore, to use multi-channel data for analyses, which would allow us to grasp a greater understanding of what the students were doing during the learning session.

Furthermore, there were limitations to the data mining techniques that were applied in this study. In the MetaTutor Study, SRL processes are measured as events, which are initiated when students click on the SRL palette, and are ended when another process begins. Thus, the system makes the assumption that: an SRL event had ended, when, in fact, it has not; or that an SRL process had ended after the student had actually completed engaging in the strategy. This can result in either two events overlapping, or a timing of an event being inaccurately captured in the log file. Furthermore, we did not measure how the sequences were influential in learning. For example, data mining analysis informed us of the most commonly used quintet sequences for HPK and LPK groups; however we did not measure the impact these sequences made on students' post-test scores; thus we cannot determine the influence of the sequences on learning, and so we could not determine if these sequences were effective for learning.

Due to the abovementioned limitations to this study, there are future directions, which we can incorporate into planning subsequent studies, which examine the effect of prior knowledge on learning in multi-agent, hypermedia-learning environments.

Future Directions and Educational Implications

The results obtained from this study lead to many future directions for analyses and design of hypermedia environments, which stress the importance of assessing and accommodating to prior knowledge groups. Future analyses on prior knowledge will account for the limitations in this study. Theoretically, studies will improve the assessment of prior knowledge, by determining prior knowledge levels based on more reliable methods, such as evaluating previous school test scores, or by assessing the students during a session prior to the learning session. Such methods would be a better measure of prior knowledge, compared to a median split based on a multiple choice test where students who do not know the responses have a chance of getting the answers correct. Methodologically, studies on prior knowledge will include more participants, which will be better obtained if a median split is not performed, and which will allow for better generalization of obtained results. Additionally, future studies will include participants in the same experimental condition. The pedagogical agents do not prompt participants in the control condition, and so a measure of students' use of SRL processes in the control condition (and not the prompt and feedback condition) will be solely based on what the students initiate during learning. It will be beneficial, therefore, to continue to collect data with participants in both conditions, and to analyze the results by condition and then compare these obtained results in addition to results obtained from participants in each experimental condition. Analytically, future studies will incorporate the analysis of multi-channel data, such that results will not be determined solely based on log-file data, which limited us to analyzing frequency data, and will include eye-tracking, physiological, audio, and video data, which will allow us to grasp a greater understanding of what students are doing at each moment of the learning session (Azevedo et al., 2013). Furthermore, future studies will expand on the data mining analyses, which were performed in this study, and which will incorporate the use of quantitative data and will assess student performance throughout and following the learning session. Lastly, it was previously mentioned that the results obtained from this study can have important implications for designing and testing the effect of prior knowledge on learning outcomes. Future studies should, therefore, examine the effect of learning and many other variables, which can be influenced by students' prior knowledge.

Moreover, the data mining analysis were an insightful preliminary analysis of how we can detect patterns of use of cognitive and metacognitive SRL strategies, and how these patterns may be different between prior knowledge groups (Bouchet, Harley, Trevors, & Azevedo, 2013; Bouchet, Kinnebrew, Biswas, & Azevedo, 2012). Therefore, future studies can further assess these patterns, while including additional variables, which are measured during the learning session. More specifically, it can be beneficial to examine the influence of the most and least commonly used strategies on learning, such that we can include post-test scores in our analysis, and use additional data mining techniques to examine the effect of the SRL sequences on learning and performance (e.g., Kinnebrew et al., 2013).

Additionally, this study mentioned the concept of cognitive load in generating hypotheses, although this construct was not measured in this study, It can be proposed that LPK students may use less SRL strategies based on cognitive overload. Possible reasons regarding this issue could assume that these students feel they have to compensate for their lack of knowledge on the content by engaging in many SRL cognitive strategies. Future studies could, therefore, assess cognitive load in students during learning, including assessing the correlation between evidence of cognitive load (e.g., pupil dilation) in students and the number of SRL processes these students engage in, in order to test the proposed hypothesis, along with many others. We would address such issues as: (1) how would we define cognitive load; (2) how would we measure cognitive load; (3) when does cognitive overload appear in students; and (4) how can we design agents to determine how and when to help students who are experiencing cognitive overload during learning. Such analyses will require real-time analysis of student performance, which we hope to make available in the newer versions of MetaTutor.

Based on the findings from this study, we can design multi-agent systems with pedagogical agents, who can be designed to adapt their decision-making for students based on the students' levels of prior knowledge. For example, the results demonstrated that students with HPK engaged in some sub goals more frequently than LPK students, and LPK students engaged in other sub goals more frequently than HPK students. Pedagogical agents can be designed to assign sub goals to students, based on their prior knowledge of the content. Furthermore, agents can be designed to monitor student performance in real-time, which will allow for agents to provide scaffolding to students at times where students appear to be having difficulties, or if students might be engaging in maladaptive SRL strategies (Azevedo & Feyzi-Behnagh, 2010). It can be beneficial to design pedagogical agents who are capable of adapting to students' individual differences, which will allow for the most optimal learning environment and can cater to each student's individual learning needs.

It was noted in the limitations section that SRL strategies, which are considered events in MetaTutor, may overlap with other events (i.e., may occur in parallel), or may be over-estimated in the time of use (e.g., if a student takes notes and then does an MPTG, we are only informed of the time when the student starts taking notes and starts engaging in the MPTG; and so it seems as though the participant took notes from the start time until he or she started the MPTG, even if he or she completed taking notes prior to engaging in the MPTG). Future systems, therefore, can work to more accurately determine the time of each SRL strategy, which will give a better measure of the time students spend engaging in SRL processes; and which will help us to better differentiate between prior knowledge groups. We can benefit from creating more enhanced multi-agent hypermedia-learning environments for students so we can cater to each student's needs, such as considering levels of prior knowledge, which can be more effective in teaching students to become better self-regulators of their learning.

References

Aleven, V., Roll, I., McLaren, B. M., Koedinger, K. R. (2010). Automated, unobtrusive, action-by-action assessment of self-regulated learning with an intelligent tutoring system. *Educational Psychologist, 45*(4), 224-233.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive Tutors: Lessons Learned. *The Journal of the Learning Sciences, 4*(2), 167–207.

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist*

(Special Issue on Computers as Metacognitive Tools for Enhancing Student Learning), 40(4), 199–209.

Azevedo, R. (2007). Understanding the complex nature of self-regulatory processes in learning with computer-based learning environments: an introduction. *Metacognition and Learning, 2*, 57-65.

Azevedo, R. (2009). Theoretical, methodological, and analytical challenges in the research on metacognition and self-regulation: A commentary. *Metacognition & Learning, 4*(1), 87–95.

Azevedo, R., & Aleven, V. (2013). Metacognition and learning technologies: An overview of the current interdisciplinary research. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp.1-16). Amsterdam, The Netherlands: Springer.

Azevedo, R., & Feyzi-Behnagh, R. (2010). Dysregulated learning with advanced learning technologies. In R. Pirrone, R. Azevedo, & G. Biswas (Eds.), *Proceedings of the AAAI Fall Symposium on Cognitive and Metacognitive Educational Systems* (pp. 5-10). Menlo Park, CA: Association for the Advancement of Artificial Intelligence (AAAI) Press.

Azevedo, R., Feyzi-Behnagh, R., Duffy, M., Harley, J., & Trevors, G. (2012). Metacognition and self-regulated learning in student-centered learning environments. In D. Jonassen and S. Land (Eds), *Theoretical foundations of learning environments*. NY: Routledge.

Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, R., Bouchet, F., & Landis, R. (2013) Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 427-449). Amsterdam, The Netherlands: Springer.

Azevedo, R., Landis, R.S., Feyzi-Behnagh, R., Duffy, M., Trevors, G., Harley, J., Bouchet, F., Burlison, J., Taub, M., Pacampara, N., Yeasin, M., Rahman, A.K.M.M., Tanveer, M.I., & Hossain, G. (2012a). The effectiveness of pedagogical agents' prompting and feedback in facilitating co-adapted learning with MetaTutor. In S. A. Cerri, W. J. Clancey, G. Papadourakis, & K. Panourgia (Eds.), *Proceedings of the 11th International Conference of Intelligent Tutoring Systems* (pp. 212-221). Amsterdam, The Netherlands: Amsterdam.

Berkhin, P. (2006). A survey of clustering data mining technques. In J. Kogan, C. Nicholas, & M. Teboulle (Eds.), *Grouping Multidimensional Data:* (pp. 25-71). Heidelberg, Berlin: Springer.

Biswas, G., Jeong, H., Kinnebrew, J., Sulcer, B., & Roscoe, R. (2010). Measuring Self-regulated Learning Skills through Social Interactions in a Teachable Agent Environment. *Research and Practice in Technology-Enhanced Learning, 5*(2), 123-152.

Bouchet, F., Harley, J., Trevors, G., & Azevedo, R. (2013). Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning. *Journal of Educational Data Mining, 5* (1), 104-146.

Bouchet, F., Kinnebrew, J., Biswas, G., & Azevedo, R. (2012). Identifying students' characteristic learning behaviors in an intelligent tutoring system fostering self-regulated learning. In K. Yacef, O. Zaiane, A. Hershkovitz, & J. Stamper (Eds.), *Proceedings of the 5th International Conference of Educational Data Mining* (pp. 65-72). Amsterdam, The Netherlands: Amsterdam.

D'Mello, S.K., Chauncey Strain, A., Olney, A., & Graesser, A. (2013). Affect, meta-affect, and affect regulation during complex learning. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 669-681). Amsterdam, The Netherlands: Springer.

Ericsson, K., Charness, N., Feltovish, P.J., & Hoffman R. (2006). *The Cambridge handbook of expertise and expert performance*. Cambridge, MA: Cambridge University Press.

Graesser, A., Chipman, P., King, B., McDaniel, B., & D Mello, S. (2007). Emotions and learning with auto tutor. *Frontiers in artificial intelligence and applications*, *158*.

Graesser, A., & McNamara, D. (2010). Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist, 45*(4), 234-244.

Greene, J.A., & Azevedo, R. (2010). The measurement of learners' self-regulated cognitive and metacognitive processes while using computer-based learning environments. *Educational Psychologist, 45*(4), 203-209.

Hegland, M. (2001). Data mining techniques. *Acta Numerica, 10*, 313-355. doi:10.1017/S0962492901000058.

Jonassen, D. & Land, S. (Eds.)(2012). *Theoretical foundations of learning environments*. NY: Routledge.

Kinnebrew, J., Biswas, G., Sulcer, B., & Taylor, R. (2013). Investigating self-regulated learning in teachable agent environments. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 451-470). Amsterdam, The Netherlands: Springer.

Lajoie, S.P., Naismith, L., Poitras, E., Hong, Y., Cruz-Panesso, Ranellucci, J., Mamane, S., & Wiseman, J. (2013). Technology-rich tools to support self-regulated learning and performance in medicine. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 427-449). Amsterdam, The Netherlands: Springer.

Lester, J.C., Mott, B.W., Robinson, J.L., & Rowe, J.P. (2013). Supporting self-regulated science learning in narrative-centered learning environments. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 471-483). Amsterdam, The Netherlands: Springer.

Mathan, S.A., & Koedinger, K.R. (2005). Fostering the intelligent novice: Learning from errors with metacognitive tutoring. *Educational Psychology, 40*(4), 257-265.

Moos, D.C., & Azevedo, R. (2008). Self-regulated learning with hypermedia. The role of prior domain knowledge. *Contemporary Educational Psychology, 33*, 270-298.

Moos, D.C., & Azevedo, R. (2009). Self-efficacy and prior domain knowledge: to what extent does monitoring mediate their relationship with hypermedia learning? *Metacognition and Learning, 4*, 197-216.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent Developments. *Educational Psychology, 38*(1), 1-4.

Pekrun, R., Goetz, T., Frenzel, A., Barchfeld, P., & Pery, R. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, 36, 36-48.

Pujari, A. K. (2001). *Data mining techniques*. India: Universities Press Private Limited.

Sawyer, K. (2006). *The Cambridge handbook of the Learning Sciences*. Cambridge, MA: Cambridge University Press.

Scheiter, K., Gerjets, P., Vollmann, B., & Catrambone, R. (2009). The impact of learner characteristics on information utilization strategies, cognitive load experienced, and performance in hypermedia learning. *Learning and Instruction, 19*, 387-401.

Shapiro , A. (1998). The relationship between prior knowledge and interactive organizers during hypermedia-aided learning. *Journal of Educational Computing Research, 20*, 143-163.

Shapiro, A. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal, 41*, 159-189.

Shapiro, A. (2008). Hypermedia design as learner scaffolding. *Educational Technology Research and Development, 56*, 29-44.

Taub, M., Azevedo, R., Bouchet, F., Feyzi-Behnagh, R., & Harley, J. (2013, August). *Can prior knowledge adequately predict the use of metacognitive processes* *during hypermedia learning*. Paper submitted for presention at the 18th biennial meeting of the European Association for Research on Learning and Instruction (EARLI), Munich, Germany.

Van Seters, J.R., Ossevoort, M.A., Tramper, J., & Goedhart, M.J.(2012). The influence of student characteristics on the use of adaptive e-learning material. *Computers & Education, 58*, 942-952.

Winne, P. H. , & Hadwin, A.F. (1998). Studying as self-regulated learning. In D.J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277-304). Hillsdale, NJ: Erlbaum.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531-566). Orlando, FL: Academic Press.

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., 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.

Woolf, B. (2009). *Building intelligent interactive tutors*. Amsterdam, The Netherlands: Morgan Kaufmann.

Zimmerman, B. & Schunk, D. (Eds.) (2011). *Handbook of self-regulation of learning and performance*. New York: Routledge.

Appendix A: Sample Pretest Questions on the Human Circulatory System

1. Nicotine causes arteries to constrict. What might happen if Mr. Smith, whose coronary arteries are partially blocked by plaque, smokes cigarettes?
   a. The nicotine might affect his breathing.
   b. The diameter of the arteries might increase in response to the nicotine
   c. The arteries might completely constrict and lead to a heart attack
   d. The nicotine might enlarge the arteries and repair the damage

2. What is the effect of the clotting process?
   a. Undigested food is eliminated
   b. Antibodies are released to fight infection
   c. Bleeding is stopped and damaged blood vessels are repaired
   d. Waste products are picked up from the body

3. The American Heart Association recommends that about 25% of a person's daily calories should come from fat. Mr. Spencer's diet is 40% fat. What situation may result from this?
   a. Increased blood clotting
   b. The increase of plaque buildup in his arteries
   c. Poor antibody production
   d. An increase in the size of his heart

4. What might happen in a disease when alveoli are stiff and not very flexible?
   a. It might be easier to send fats to the liver
   b. It might be more difficult for gas exchange to occur
   c. They might not be connected to the bronchial tubes
   d. They might not be surrounded by capillaries

5. What are the tiny air sacs that are found at the end of the branches of the bronchial tubes?
   a. Capillaries
   b. Lungs
   c. Alveoli
   d. Glands

6. Why is the surface area of the alveoli so large?
   a. To make the alveoli more extensive than the skin
   b. To aid in digestion
   c. To speed up diffusion of gases in the lungs
   d. To enable the alveoli to inflate with inhaled air

Appendix B: Experimental Set-up for Collecting Data with MetaTutor

Appendix C: Tables of Numerical Data for Research Question 5

Table 4.

Number of SRL Processes vs. Time Spent on SRL Processes for Path of Blood Flow Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.333	0.376
1 – HPK	0.0	0.0
1 – HPK	0.1	0.004
1 – HPK	0.133	0.013
1 – HPK	0.1	0.017
1 – HPK	0.267	0.347
1 – HPK	0.1	0.004
1 – HPK	0.4	0.141
2 – LPK	0.0	0.0
2 – LPK	0.067	0.0002
2 – LPK	0.067	0.010
2 – LPK	1.0	0.838
2 – LPK	0.167	0.150
2 – LPK	0.033	0.0

Table 5.

Number of SRL Processes vs. Time Spent on SRL Processes for Heartbeat Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.16	0.026
1 – HPK	0.0	0.0
1 – HPK	1.0	1.0
1 – HPK	0.08	0.018
1 – HPK	0.2	0.003
1 – HPK	0.08	0.004
1 – HPK	0.16	0.092
1 – HPK	0.04	0.007
2 – LPK	0.0	0.0
2 – LPK	0.053	0.002
2 – LPK	1.0	0.175
2 – LPK	0.053	0.0
2 – LPK	0.053	0.002
2 – LPK	0.158	0.021

Table 6.

Number of SRL Processes vs. Time Spent on SRL Processes for Heart Components Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.0	0.0
1 – HPK	0.429	0.617
1 – HPK	0.679	0.761
1 – HPK	0.393	0.666
1 – HPK	0.179	0.041
1 – HPK	0.321	0.389
1 – HPK	0.25	0.205
2 – LPK	0.0	0.0
2 – LPK	0.036	0.00009
2 – LPK	0.107	0.005
2 – LPK	1.0	0.461
2 – LPK	0.143	0.041
2 – LPK	0.393	0.145
2 – LPK	0.571	0.012
2 – LPK	0.25	0.137

Table 7.

Number of SRL Processes vs. Time Spent on SRL Processes for Blood Vessels Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.0	0.0
1 – HPK	0.053	0.006
1 – HPK	0.105	0.036
1 – HPK	0.053	0.005
1 – HPK	0.053	0.0
1 – HPK	0.263	0.275
2 – LPK	1.0	1.0
2 – LPK	0.0	0.0
2 – LPK	0.632	0.219
2 – LPK	0.158	0.007
2 – LPK	0.947	0.961
2 – LPK	0.158	0.005
2 – LPK	0.474	0.142

Table 8.

Number of SRL Processes vs. Time Spent on SRL Processes for Blood Components Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.0	0.0
1 – HPK	0.444	0.345
1 – HPK	0.148	0.007
1 – HPK	0.222	0.353
1 – HPK	0.556	0.131
1 – HPK	0.519	1.0
1 – HPK	0.259	0.113
1 – HPK	1.0	0.797
1 – HPK	0.296	0.051
2 – LPK	0.0	0.0
2 – LPK	1.0	0.016
2 – LPK	0.889	0.043
2 – LPK	0.333	0.002
2 – LPK	0.556	0.009
2 – LPK	0.333	0.002
2 – LPK	1.0	0.056

Table 9.

Number of SRL Processes vs. Time Spent on SRL Processes for Purposes of the Circulatory System Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.0	0.0
1 – HPK	0.5	0.345
1 – HPK	0.75	0.434
1 – HPK	0.35	0.189
1 – HPK	0.2	0.005
1 – HPK	0.85	0.580
2 – LPK	0.0	0.0
2 – LPK	0.7	0.611
2 – LPK	0.05	0.004
2 – LPK	1.0	0.522
2 – LPK	0.1	0.006
2 – LPK	0.4	0.075
2 – LPK	0.3	0.041
2 – LPK	0.15	0.004

Table 10.

Number of SRL Processes vs. Time Spent on SRL Processes for Malfunctions of the Circulatory System Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.0	0.0
1 – HPK	0.333	0.054
1 – HPK	0.467	0.072
1 – HPK	0.533	0.365
1 – HPK	0.733	0.070
1 – HPK	1.0	1.0
1 – HPK	0.467	0.294
1 – HPK	0.733	0.646
1 – HPK	0.067	0.0
1 – HPK	0.067	0.011
1 – HPK	0.067	0.009
2 – LPK	0.0	0.0
2 – LPK	0.111	0.0
2 – LPK	0.222	0.005
2 – LPK	0.444	0.002
2 – LPK	1.0	0.196

Appendix D: Tables of Numerical Data for Research Question 6

Table 11.

Number of SRL Processes vs. Page Relevancy for Path of Blood Flow Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.333	1.0
1 – HPK	0.0	0.0
1 – HPK	0.1	0.490
1 – HPK	0.133	0.312
1 – HPK	0.1	0.315
1 – HPK	0.267	0.471
1 – HPK	0.1	0.0
1 – HPK	0.4	0.435
2 – LPK	0.0	0.0
2 – LPK	0.067	0.222
2 – LPK	0.067	0.462
2 – LPK	1.0	0.333
2 – LPK	0.167	0.290
2 – LPK	0.033	0.286

Table 12.

Number of SRL Processes vs. Page Relevancy for Heartbeat Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.16	1.0
1 – HPK	0.0	0.0
1 – HPK	1.0	0.057
1 – HPK	0.08	0.8
1 – HPK	0.2	0.667
1 – HPK	0.08	0.357
1 – HPK	0.16	0.667
1 – HPK	0.04	0.8
2 – LPK	0.0	0.0
2 – LPK	0.053	0.222
2 – LPK	1.0	0.154
2 – LPK	0.053	0.118
2 – LPK	0.053	0.182
2 – LPK	0.158	1.0
2 – LPK	0.053	0.111

Table 13.

Number of SRL Processes vs. Page Relevancy for Heart Components Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.0	0.0
1 – HPK	0.429	0.5
1 – HPK	0.679	0.357
1 – HPK	0.393	1.0
1 – HPK	0.179	0.308
1 – HPK	0.321	0.455
1 – HPK	0.25	0.875
2 – LPK	0.0	0.0
2 – LPK	0.036	0.524
2 – LPK	0.107	0.462
2 – LPK	1.0	0.8
2 – LPK	0.143	0.571
2 – LPK	0.393	0.833
2 – LPK	0.571	0.526
2 – LPK	0.25	0.857

Table 14.

Number of SRL Processes vs. Page Relevancy for Blood Vessels Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.0	0.0
1 – HPK	0.053	1.0
1 – HPK	0.105	1.0
1 – HPK	0.053	0.611
1 – HPK	0.053	0.684
1 – HPK	0.263	0.528
2 – LPK	1.0	0.846
2 – LPK	0.0	0.0
2 – LPK	0.632	0.833
2 – LPK	0.158	0.5
2 – LPK	0.947	0.75
2 – LPK	0.158	1.0
2 – LPK	0.474	0.714

Table 15.

Number of SRL Processes vs. Page Relevancy for Blood Components Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.0	0.0
1 – HPK	0.444	0.417
1 – HPK	0.148	0.556
1 – HPK	0.222	1.0
1 – HPK	0.556	0.5
1 – HPK	0.519	0.471
1 – HPK	0.259	0.417
1 – HPK	1.0	0.304
1 – HPK	0.296	0.313
2 – LPK	0.0	0.0
2 – LPK	1.0	0.395
2 – LPK	0.889	1.0
2 – LPK	0.333	0.539
2 – LPK	0.556	0.0
2 – LPK	0.333	1.0
2 – LPK	1.0	0.818

Table 16.

Number of SRL Processes vs. Page Relevancy for Purposes of the Circulatory System Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.0	0.0
1 – HPK	0.5	1.0
1 – HPK	0.75	0.595
1 – HPK	0.35	1.0
1 – HPK	0.2	0.469
1 – HPK	0.85	0.625
2 – LPK	0.0	0.0
2 – LPK	0.7	0.571
2 – LPK	0.05	0.704
2 – LPK	1.0	0.667
2 – LPK	0.1	0.654
2 – LPK	0.4	0.591
2 – LPK	0.3	0.679
2 – LPK	0.15	0.438

Table 17.

Number of SRL Processes vs. Page Relevancy for Malfunctions of the Circulatory System Sub Goal, based Prior Knowledge Group

Group	X-Coordinate	Y-Coordinate
1 – HPK	0.0	0.0
1 – HPK	0.333	1.0
1 – HPK	0.467	0.273
1 – HPK	0.533	0.071
1 – HPK	0.733	0.636
1 – HPK	1.0	1.0
1 – HPK	0.467	0.583
1 – HPK	0.733	0.444
1 – HPK	0.067	0.2
1 – HPK	0.067	0.243
1 – HPK	0.067	0.438
2 – LPK	0.0	0.0
2 – LPK	0.111	0.0
2 – LPK	0.222	0.857
2 – LPK	0.444	0.7
2 – LPK	1.0	1.0