Analyzing Students' Metacognitive Strategies in Open-Ended Learning Environments

Gautam Biswas (gautam.biswas@vanderbilt.edu)
Department of EECS/ISIS, 1025 16th Ave South
Nashville, TN 37212 USA

John S. Kinnebrew (john.s.kinnebrew@ vanderbilt.edu)
Department of EECS/ISIS, 1025 16th Ave South
Nashville, TN 37212 USA

James R. Segedy (james.segedy@ vanderbilt.edu)
Department of EECS/ISIS, 1025 16th Ave South
Nashville, TN 37212 USA

Abstract

Novices often lack metacognition and self-regulation skills that are important for effective learning. Betty's Brain, an open-ended computer-based learning environment helps students practice and develop metacognitive strategies as they learn science topics. We extend previous work on sequence mining methods to discover students' frequently-used behavior patterns from their activity sequences. Our results show that it is possible to interpret aspects of students' learning strategies and their effectiveness by taking into account the context of their activities in the system.

Keywords: open-ended learning environments, metacognition, measuring metacognition, scaffolding, sequence mining.

Introduction

Cognitive scientists have established that metacognition and self-regulation are essential for developing effective learning strategies in the classroom and beyond (Bransford et al, 2000; Zimmerman, 2001). However, novice learners often have ineffective self-regulation profiles, which may be attributed to their lacking the well-organized domain knowledge structures of experts. This affects their ability to break down their learning and problem solving into distinct task understanding, planning and solution generation, monitoring and evaluation phases, leading them to use suboptimal learning and problem solving strategies (Chi et al, 1988; VanLehn, 1996).

Our research group has developed Betty's Brain, an open-ended learning environment (OELE), to study how students develop metacognitive strategies that include constructing information and monitoring as they learn science topics (Leelawong and Biswas, 2008). Our approach utilizes trace methodologies derived from students' actions and activity patterns in the environment to infer aspects of their metacognitive abilities (Aleven et al, 2006; Azevedo, et al., 2012; Hadwin et al, 2007). This is based on a *metacognition as events* hypothesis, which theorizes that the use of metacognitive strategies manifests as continually unfolding events that can be inferred from learners' behaviors.

In this paper, we extend our previous work on using sequence mining methods to discover students' frequently-used behavior patterns from their activity sequences as they work in the Betty's Brain system (Kinnebrew & Biswas, 2012). In particular, we extend our techniques for analyzing students' action sequences by (i) interpreting and characterizing behavior patterns using a cognitive/metacognitive model of the task, (ii) mapping students' frequently observed cognitive and metacognitive process patterns back into their overall activity sequences, and (iii) using metrics to evaluate the effectiveness of these processes. The results in this paper represent a post hoc analysis of student behaviors, but our longer term goal is to use such results to monitor and measure students' cognitive and metacognitive processes online as they work on their learning and problem-solving tasks, and use these results to develop adaptive scaffolding mechanisms that support student learning.

Background

Metacognition is often described as being made up of two constituent parts (Flavell et al, 1985; Veenman, 2012): (1) *Metacognitive knowledge*, which is declarative and deals with the interplay between knowledge of one's abilities to perform tasks, the nature of the task, and the strategies one can employ to successfully perform the task; and (2) *Metacognitive monitoring* and regulation, which includes activities related to planning, monitoring, and evaluating one's cognitive processes in order to better regulate those processes in the future.

Researchers have established strong links between learners' metacognitive abilities and their effectiveness in executing cognitive processes. Winne (1996) characterizes cognition as dealing with knowledge of objects and operations on objects (the object level) while characterizing metacognition as the corresponding meta-level that contains information about cognitive processes. Metacognitive monitoring brings the two levels together, as it describes the process of observing one's own execution of cognitive processes at the object level and exerting control over the object level using metacognitive knowledge and strategies.

An important implication of the interplay between cognition and metacognition relates to the dependence of metacognition on cognition (Land, 2000). In other words, metacognitive knowledge may not be sufficient for achieving success in learning and problem solving, especially for learners who lack the cognitive skills and background knowledge necessary for interpreting, understanding, and organizing critical aspects of the problem under study (Bransford et al, 2000). Learners may also lack knowledge of effective strategies (e.g., the ability to extract relevant information when reading a science text), and, therefore, resort to suboptimal strategies in performing their tasks (Azevedo, 2005; Kinnebrew & Biswas, 2012). Poor self-judgment abilities result in difficulties for monitoring and evaluating one's own effectiveness and progress, which can be a significant stumbling block in selecting and implementing relevant strategies in a timely manner.

However, research studies have shown that with proper scaffolding, middle school students can improve their metacognitive awareness and develop effective metacognitive strategies (Kramarski & Mevarech, 2003). Our system, Betty's Brain is designed to help middle school students develop metacognitive knowledge and strategies as they learn about science topics. Other systems with similar goals include MetaTutor (Azevedo, et al., 2012) and Crystal Island (Rowe, et al., 2011).

Betty's Brain

Betty's Brain (Figure 1) is an open-ended learning environment (Land, 2000) that provides students with a learning context and a set of tools for pursuing authentic and complex learning tasks. Students teach a virtual agent, Betty, about science topics by constructing a causal map. The goal for students using Betty's Brain is to teach Betty a map, whose correctness is determined in relation to a hidden, expert causal map.

The students' learning and teaching tasks are organized around three activities: (1) reading hypertext resources to learn the domain material, (2) building and refining a causal map, which represents the domain material, and (3) asking Betty to take a quiz. Students explicitly teach Betty by constructing a causal map. For example, they may draw a causal link between *garbage and landfills* and *methane* to represent the relationship *garbage and landfills increase methane* (a greenhouse gas). Students can check what Betty knows by asking her questions, e.g., *if garbage and landfills decrease, what effect does it have on polar sea ice*? To answer questions, Betty uses qualitative reasoning that operates through chains of links from the source concept to the target concept (Leelawong & Biswas, 2008). The learner can further probe Betty's understanding by asking her to explain her answer. Betty illustrates her reasoning by explaining her thinking and animating her explanation by highlighting concepts and links on the map as she mentions them.

Learners can assess Betty's (and, therefore, their own) progress in two ways. After Betty answers a question, learners can ask Mr. Davis, a pedagogical agent that serves as a mentor, to evaluate the answer. Learners can also have Betty take a quiz on one or all of the sub-topics in the resources. Quiz questions are selected dynamically to reflect the current state of the student's map; questions are chosen (in proportion to the completeness of the map) for which Betty will generate correct answers. The remaining questions produce incorrect answers, and they direct the student's attention to incorrect and missing links.

After Betty takes a quiz, her results, including the causal map she used to answer the questions appear on the screen as shown in Figure 1. The quiz questions, Betty's answer, and the Mentor's assigned grade, i.e., correct, correct but incomplete, or incorrect appear on the top of the window. Clicking on a question will highlight the causal links that Betty used to answer that question. To help students keep track of correct and incorrect links, the system allows students to annotate them with a green check-mark (correct), a red X (incorrect), or a gray question mark (not sure).

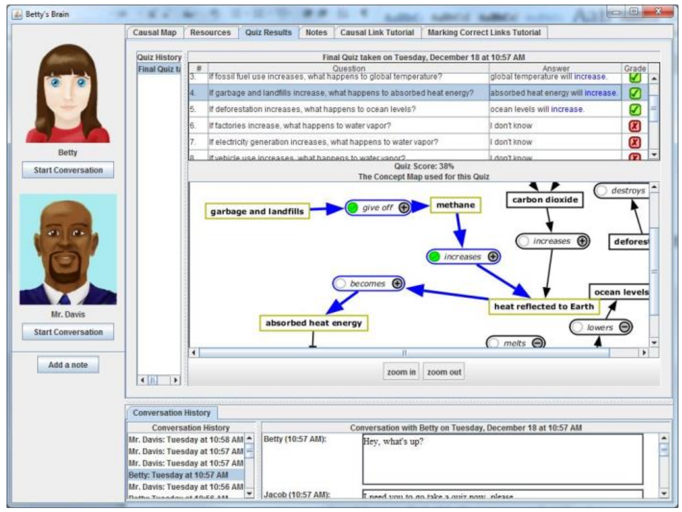


Figure 1: Betty's Brain Interface with Quiz Window

Cognitive/Metacognitive Process Model

To interpret students learning behaviors on the system, we have developed a model that takes into account the tight connection between the cognitive and metacognitive processes needed to address the learning task effectively. Overall, this model includes four primary processes that students are expected to engage in while using Betty's Brain: (1) Goal Setting & Planning, (2) Knowledge Construction (KC), (3) Monitoring (Mon), and (4) Help Seeking. In this work we focus on the KC and Mon process models.

Knowledge construction includes metacognitive strategies for (1) information seeking, i.e., determining when and how to locate needed information in the resources, and (2) information structuring, i.e., organizing one's developing understanding of the domain knowledge into structural components (e.g., causal links). In executing these metacognitive processes, learners have to apply relevant cognitive processes listed under information seeking and structuring. Seeking information, for example, requires that students to identify the causal information by reading the resources and making sense of the content. Similarly, information structuring captures the process of successfully converting the acquired information into causal links and adding them to the causal map.

Monitoring processes include (1) model assessment, i.e., assessing the correctness of all or a part of the causal model, and (2) progress recording, i.e., making explicit annotations to mark parts of the causal model as correct, which makes it easier to focus on parts of the map that need more work. Successful execution of monitoring metacognitive processes relies on students' abilities to execute cognitive processes for assessing the causal model (via questions, explanations, quizzes, and question evaluations) and recording progress (via note taking and annotating links with correctness information). The cognitive and metacognitive process model provides a framework for interpreting students learning activities and behaviors (activity sequences) on the system.

Measuring Cognition and Metacognition

We have developed a set of data mining methods for analyzing students' learning activity sequences and assessing their learning processes as they work in Betty's Brain. In addition, we have developed visualization methods for measuring how student behaviors evolve during the course of the intervention depending on the type of feedback and support that they received from the Mentor agent. In particular, we were interested in studying whether students' suboptimal behaviors were replaced by more optimal strategies as the intervention progressed.

To assess student activities with respect to our cognitive/metacognitive model, we calculate four measures: *map edit effectiveness*, *map edit support*, *monitoring effectiveness*, and *monitoring support*. Map edit effectiveness is calculated as the percentage of causal link additions, removals, and modifications that improve Betty's causal map. Map edit support is defined as the percentage of causal map edits that are *supported* by previous reading of pages in the resources that discuss the concepts connected by the manipulated causal link. Monitoring effectiveness is calculated as the percentage of quiz questions and explanations that generate specific correctness information about one or more causal links. For example, all of the links used in a quiz question whose answer is marked correct, must be correct. If the answer to a question is incorrect, at least one of the links used in the answer must be incorrect. Finally, monitoring support is defined as the percentage of causal link annotations that are supported by previous quiz questions and explanations. For support metrics, a further constraint is added: an action can only support another action if both actions occur within the same time window, and we calculated support for a ten minute time window.

The information for calculating the measures and deriving student behavior using sequence mining is extracted from log files. For example, if a student accesses a page in the resources, this is logged as a Read action that includes additional information, *e.g.*, the page accessed. In this work, students' activity sequences contain six categories of actions: (1) Read, (2) Link Edit, (3) Query, (4) Quiz, (5) Explanation, and (6) Link Annotation. Actions were further distinguished by context details, such as the correctness of a link edit. Sequence mining techniques are applied to discover frequent behavior patterns for students in a given group (Kinnebrew, et al., 2013; Kinnebrew & Biswas, 2012). Students' use of metacognitive processes was determined by interpreting the patterns using the cognitive and metacognitive model.

Method

The present analysis used data from a recent classroom study with Betty's Brain in which students learned about the greenhouse effect and climate change. The study tested the effectiveness of two support modules designed to scaffold students' understanding of cognitive and metacognitive processes important for success in Betty's Brain. The knowledge construction (KC) module provided support on how to identify causal relations in the resources, and the monitoring (Mon) support module helped students understand how to use Betty's quizzes to identify correct and incorrect causal links on the causal map. Participants were divided into three treatment groups. The KC group (KC-G) used a version of Betty's Brain that included the KC support module and a causal link tutorial that they could access at any time during learning. The tutorial allowed students to practice identifying causal relations in short text passages. The Mon group (Mon-G) used a version of Betty's Brain that included the Mon support module and a marking links correct tutorial that they could access at any time during learning. The tutorial presented practice problems in which students used the results of graded quiz questions and the causal map used to answer those questions to select the links that could be marked as correct. Finally, the control group (Con-G) used a version of Betty's Brain that included neither the tutorials nor the support modules.

The KC module was activated when three out of a student's last five map edits were incorrect, at which point Mr. Davis would begin suggesting strategies for identifying causal links during reading. Should students continue to make incorrect map edits despite this feedback, the KC module activated a second tier of support: *guided practice*. During guided practice, students were moved to the causal link tutorial where they read short text passages and expressed the primary idea in the passage as a causal relation. When they worked on the tutorial, students were not permitted to access any other portion of the program. Students completed the tutorial session once they solved five problems correctly without making a mistake.

The Mon module was activated after the third time students did not use evidence from quizzes and explanations to annotate links on their map. At this time, Mr. Davis began suggesting strategies for using quizzes and explanations to identify and keep track of which links were correct. Additionally, Mr. Davis discouraged students from annotating links as being correct without using the suggested strategies. Should students continue to use quizzes and explanations without annotating links correctly, the Mon module provided students with guided practice. Like the KC tutorial, students had to complete five problems correctly on the first try to complete the tutorial session.

Seventy-three seventh grade students from four middle Tennessee science classrooms, taught by the same teacher, participated in the study. Because use of Betty's Brain relies on students' ability to independently read and understand the resources, the system is not suited to students with limited English proficiency or cognitive-behavioral problems. Therefore, data from English as a Second Language (ESL) and special education students were not analyzed. Similarly, we excluded the data of students who missed more than two class periods of work on the system. Our experimental analysis used data collected from fifty-two students who participated in the study.

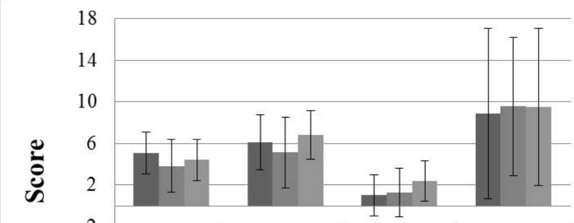
Learning was assessed using a pre-post test design. Each written test consisted of five questions that asked students to consider a given scenario and explain its causal impact on climate change. Scoring was based on the causal relations that students used to explain their answers to the questions, which were then compared to the chain of causal relations used to derive the answer from the expert map. One point was awarded for each causal relationship in the student's answer that came from or was closely related to an expert causal link. The maximum combined score for the five questions was 16. Two coders independently scored a subset of the pre- and post-tests with at least 85% agreement, at which point the coders split the remaining tests and individually coded the answers and computed the scores.

Performance on the system was assessed by calculating a score for the causal map that students created while teaching Betty. This score was computed as the number of correct links (the links in the student's map that appeared in the expert map) minus the number of incorrect links in the student's final map. We also used the log data collected from the system to derive students' behavior patterns, interpret them using our cognitive/metacognitive model, and study the temporal evolution of the observed KC and Mon strategies over the period of the intervention.

Study duration was 9 school days. During the first 60 minute class period, students completed the pre-test. During the second and third class periods, researchers introduced students to causal modeling and reasoning with causal models, and how to identify causal relations in text passages. During this time, students completed paper-and-pencil group exercises involving causal reasoning and identifying causal relations. During the fourth class period, students were provided with hands-on system training by the researchers. Students then spent four class periods using their respective versions of Betty's Brain with minimal intervention by the teachers and the researchers. On the ninth day, students completed the post-test.

Results

Figure 2 presents the overall learning and performance results for each condition in the intervention. Repeated measures ANOVA performed on the data revealed a significant effect of time on test scores (F=28.66, p <0.001). Pairwise comparison of the three groups revealed that the Mon-G had marginally better learning gains than KC-G, which had better learning gains than the Con-G group. The Mon-G learning gains were significantly better than the Con-G gains at the 0.1 significance level (p < .075), indicating the two interventions may have resulted in better understanding of the science content. The small sample size and the large variations in performance within groups made it difficult to achieve statistical significance in these results. However, one positive aspect of this finding is that while students in the Mon-G and KC-G spent an average of 10% and 17% of their time in guided practice, respectively, they learned, on average, just as much, if not more, than the Con-G students.



	Pretest	Posttest	Gain	Map Score (max = 25)
	Climate Change Short Answer (max = 16)			
Con-G	5.07 (2.03)	6.10 (2.64)	1.03 (1.99)	8.87 (8.20)
KC-G	3.58 (2.54)	5.13 (3.37)	1.28 (2.33)	9.55 (6.64)
Mon-G	4.41 (1.97	6.82 (2.33)	2.41 (1.92)	9.53 (7.55)

Figure 2: Pre-post Test Results (mean (std dev)) and Final Map Score

To assess students' overall behaviors, we calculated the effectiveness and support measures, which are illustrated in Table 1. The KC-G students had the highest scores on both map editing effectiveness and support, suggesting that the KC feedback did help students more effectively and systematically read and construct their causal maps (however, only the map edit support showed a statistically significant difference, KC-G > Con-G, p = 0.02, and the map edit effectiveness illustrated a trend, KC-G > Con-G, p = 0.08). However, the monitoring support did not help the Mon-G students do better than the other two groups for monitoring effectiveness or support. The Mon-G students did have the highest monitoring effectiveness, but it was not statistically significant. Further, the Con-G students had the monitoring support average (p < 0.10, when comparing with other groups). It is not clear why the Mon or KC support and tutorials resulted in students performing less supported monitoring activities tan the Con-G students.

Table 1: Effectiveness & Support Measures ((mean (std dev)) by Group

Measure	Con-G	KC-G	Mon-G
Map edit effectiveness	0.46 (0.13)	0.52 (0.07)	0.5 (0.12)
Map edit support	0.43 (0.25)	0.64 (0.19)	0.55 (0.23)
Monitoring effectiveness	0.3 (0.22)	0.32 (0.21)	0.4 (0.20)
Monitoring support	0.61 (0.30)	0.32 (0.4)	0.33 (0.32)

In order to investigate student learning behavior in more detail, we employed sequence mining analyses to identify 143 different action patterns that were observed in the majority of students. Table 2 lists the 10 most frequent patterns that employed at least two actions and could be interpreted as a metacognitive strategy in our cognitive/metacognitive model. Each pattern is defined by two or more primary actions, and each action is qualified by one or more attributes. For example, a [Read] → [Add correct link, relevant to recent actions] pattern describes a KC behavior, where the student added a correct causal link to the map after a [Read] action where the student read a page that discussed the added link. In contrast, the action labeled [Read] → [Add incorrect link, relevant to recent actions] implies the student added an incorrect link even after reading a page that contained information about the link. The → symbol implies that the action to the left of the arrow preceded the action to the right of the arrow.

The average frequency represents the average number of times students used a particular behavior pattern when they worked on the system. These numbers are broken down for the three conditions. The last column represents our interpretation of the type of strategy a particular behavior represents. In this study, the strategy corresponding to a behavior was determined by the category of the cognitive process (KC or Mon) implied by the individual actions that made up the behavior Therefore, some behaviors, e.g., pattern #3: [Quiz] → [Remove incorrect link], span KC and Mon (KC+Mon) strategies.

The frequency numbers indicate that for almost all of the top 10 behaviors the CON-G showed a higher frequency of use than the two experimental groups. This may be partly attributed to the time the KC-G and Mon-G groups spent in tutorials, therefore reducing the amount of time they spent on the map building task. However, an equally likely reason may be that the CON-G students used more trial-and-error approaches, spending less time editing and checking the correctness of their maps in a systematic way. This is further supported by looking at the highest average frequency behaviors for each of the groups. The top five behavior strategies for the Mon-G students are primarily Mon or KC+Mon related (patterns 1, 3, 5, 7, and 9), involving quizzes, map editing, and explanations. KC-G students, on the other hand, more often employed KC strategies related to adding and removing links along with a couple of strategies that combine KC and Mon activities. The Con-G students seem to have employed KC and Mon strategies in about equal numbers, but they were less effective in using these strategies.

An interesting strategy is pattern #10: [Add incorrect link (AIL)] → [Quiz (Q)] → [Remove incorrect link (RIL)]. This may represent a strategy where a student first adds a link (which happens to be incorrect) and then takes a quiz to determine if the quiz score changes. Depending on the outcome (in this case, the score likely decreased), the student determines that the link added was incorrect, and, therefore, removes it. This represents a trial-and-error strategy. While students in all three groups used this strategy, the Mon-G group used it with lower frequency than the other two groups, and this may be attributable to the effectiveness of the Monitoring scaffolding. To study this pattern further we developed two measures: (1) a measure of *cohesiveness* of the pattern, i.e., in what percentage of the AIL → Q → RIL patterns was the delete action supported by the quiz result; and (2) a *support* measure, i.e., in what percentage of the AIL → Q → RIL patterns was the addition of the link supported by recent actions. The MON group had higher cohesiveness (41.9 to 38.0 and 37.3 for the CON and KC groups) and support (27.7 to 20.3 and 187.7 for the CON and KC groups) measures, implying that they used this pattern in a more systematic way than the other two groups.

Discussion and Conclusions

The results presented in the previous section provide evidence that a combination of theory-driven measures and data-driven mining techniques can be successfully employed to produce a more complete description of the metacognitive strategies use in their learning and problem-solving tasks. In our work on investigating cognitive and metacognitive processes in Betty's Brain, we had to carefully instrument the system to collect rich data on the students' activities and the context associated with those activities. Post hoc mining and analysis of this data revealed a number of interesting results. Perhaps most important, the results show (i) that it is possible to infer aspects of students' use of strategies as they learn through these data mining and analysis techniques combined with a cognitive/metacognitive model of the task, and (ii) that tracking student performance and related context information with respect to their activities allows us to better characterize these strategies as suboptimal versus optimal.

Our analyses in this study focused on students' knowledge construction and monitoring strategies. Knowledge construction strategies include seeking out information, thinking deeply about the material to develop a sufficient understanding to apply it to model building and problem solving tasks. In particular, information structuring strategies in Betty's Brain help students with their map-building activities, which include understanding the structure of the causal model, the ability to construct it in parts, the ability to add links correctly to an existing structure, and also the ability to reason (e.g., answer questions, formulate hypotheses) with the evolving structure. The primary monitoring strategies relate to determining when and how to check the correctness of the current causal map, and then, in more detail, using the quiz (assessment) results to determine the correctness of individual links, and what parts of the map are incomplete or still need work.

In summary, the analysis presented in this paper successfully employed our metacognition measurement framework to evaluate the effects of scaffolding support for metacognitive and cognitive processes important for success in Betty's Brain. In particular, we applied these analyses to a comparison of different versions of Betty's Brain, a version that pro-

Table 2: Comparison of Pattern Frequencies across Conditions

Rank	Pattern	Avg. Frequency			Model Category
		CON	KC	MON	
1	[Add incorrect link] → [Quiz]	11.20	7.35	8.24	KC+Mon
2	[Add incorrect link] → [Remove incorrect link]	6.00	12.65	3.71	KC
3	[Quiz] → [Remove incorrect link]	7.87	6.10	6.29	KC+Mon
4	[Add concept] → [Add correct link]	7.53	6.75	4.94	KC
5	[Quiz] → [Explanation]	8.40	3.80	5.35	Mon
6	[Remove incorrect link] → [Add incorrect link]	4.53	9.20	3.41	KC
7	[Add correct link] → [Quiz]	5.87	4.05	5.06	KC+Mon
8	[Remove incorrect link] → [Quiz]	5.93	4.45	4.12	KC+Mon
9	[Explanation] → [Explanation]	5.67	2.95	4.88	Mon
10	[Add incorrect link] → [Quiz] → [Remove incorrect link]	5.20	4.40	3.88	KC+Mon

vided very little scaffolding and no guided practice versus two experimental conditions: one that provided KC scaffolds and a second that provided Mon scaffolds. Overall, the interventions produced changes in student behavior that were consistent with the provided scaffolding, implying that these metacognitive strategies can be taught and supported for middle school students in computer-based learning environments.

An interesting implication of this work is that monitoring strategies can be key to better learning performance, and better monitoring strategies may provide the catalyst for developing more effective knowledge construction, i.e., information seeking and information structuring strategies. The results presented in this paper are promising, but further analysis and more systematic experiments will have to be conducted to achieve conclusive results.

Future work will involve refining the methods presented in this paper in order to allow us to discover and define strategies in a more systematic way. Further, we will extend our measurement framework to more closely integrate theory-driven measures with data-driven mining for analyzing student cognition and metacognition during learning. Ultimately, we hope to find better ways of inferring students' intent (i.e., goals) from their observed behaviors and strategies while using the system.

Acknowledgments

This work has been supported by NSF-IIS Award #0904387 and IES Award #R305A120186.

References

Aleven V, McLaren B, Roll I, Koedinger K (2006) Toward meta-cognitive tutoring: A model of help seeking with a Cognitive Tutor. *International Journal of Artificial Intelligence in Education* 16(2):101-128

Azevedo, R., et al.. (2012). The Effectiveness of Pedagogical Agents' Prompting and Feedback in Facilitating Co-adapted Learning with MetaTutor. In *Intelligent Tutoring Systems* (pp. 212-221). Springer Berlin/Heidelberg.

Bransford J, Brown A, Cocking R (eds) (2000) *How people learn*. National Academy Press Washington, DC, Washington.

Chi M, Glaser R, Farr M (1988) *The nature of expertise*. Lawrence Erlbaum Associates, Inc.

Flavell J, Miller P, Miller S (1985) *Cognitive development*. Prentice-Hall Englewood Cliffs, NJ

Hadwin A, Nesbit J, Jamieson-Noel D, Code J, Winne P (2007) Examining trace data to explore self-regulated learning. *Metacognition and Learning*, 2(2):107-124.

Kinnebrew J.S., Biswas G. (2012) Identifying learning behaviors by contextualizing differential sequence mining with action features and performance evolution. In: *Proc. 5th International Conference on Educational Data Mining* (EDM 2012), Chania, Greece

Kinnebrew J.S., Loretz KM, Biswas G (2013) A contextualized, differential sequence mining method to derive students' learning behavior patterns. *Journal of Educational Data Mining*

Kramarski B., Mevarech Z. (2003) Enhancing mathematical reasoning in the classroom: The effects of cooperative learning and metacognitive training. *American Educational Research Journal* 40(1):281-310

Land S. (2000) Cognitive requirements for learning with open-ended learning environments. *Educational Technology Research and Development*, 48(3):61-78

Leelawong K, Biswas G (2008) Designing learning by teaching agents: The Betty's Brain system. *International Journal of Artificial Intelligence in Education*, 18(3):181-208.

Rowe, J., Shores, L., Mott, B., and Lester, J. (2011). Integrating learning, problem solving, and engagement in narrative-centered learning environments. *Intl Jour of Artificial Intelligence in Education*, 21(1-2):115–133.

VanLehn K (1996) Cognitive skill acquisition. *Annual review of psychology* 47(1):513-539.

Veenman M (2012) Metacognition in science education: Definitions, constituents, and their intricate relation with cognition. *Metacognition in Science Education*, pp 21-36.

Winne P (1996) A metacognitive view of individual differences in self-regulate learning. *Learning and individual differences* 8(4):327-353.

Zimmerman B (2001) Theories of self-regulated learning and academic achievement: An overview and analysis. In: Zimmerman B, Schunk D (eds) *Self-regulated learning and academic achievement: Theoretical perspectives*, Erlbaum, Mahwah, NJ, pp 1-37.