Missions with Monty: A Game-Based Learning Environment to Promote Comprehension Monitoring and Science Achievement

Rayne A. Sperling[1], John L. Nietfeld[2], Samira Syal[2] and Taylor M. Young[1]
[1]The Pennsylvania State University
[2]North Carolina State University
rsd7@psu.edu
john_nietfeld@ncsu.edu
srajesh@ncsu.edu
tmy5110@psu.edu

Abstract: While game-based learning environments (GBLE) are an increasingly popular means to engage learners, there remains limited research on the benefits of GBLEs for students' learning and metacognitive skills in classroom contexts. Missions with Monty is a GBLE developed to engage students as they learn comprehension strategies to support their scientific literacy with targeted Next Generation Science Standards content. The current missions focus on ecosystems, earth and human activity, molecules, and organisms. Students navigate through the game to solve challenges that require the application and transfer of science knowledge. One primary and unique focus of Missions with Monty is to benefit students' metacognition through comprehension monitoring. Comprehension monitoring is an essential strategy to support literacy and is measured through students' accuracy in rating their performance on comprehension items. In the game, students are explicitly taught and subsequently practice essential comprehension monitoring strategies. Additional key educational outcomes for students playing Missions with Monty include reading achievement, science knowledge, and transfer of science learning. In this study we examined 10 and 11 year old students' outcomes as they either played the game ($n = 144$) or learned the content via a computer-based learning environment (CBLE) without game elements in a comparison condition ($n = 80$) within classrooms of diverse learners in the United States. Significant gains for science learning were found across conditions, ($t(223) = 13.67$, $p <.001$). The primary research question of interest focused on whether students who played Missions with Monty were better able to monitor their understanding of science text after playing the game and therefore we compared monitoring, as measured by bias and accuracy, before and after the game play. Effective metacognition is demonstrated by low bias and high accuracy. Findings revealed that both conditions demonstrated improved bias ($t(223) = 4.44$, $p < .001$) and accuracy ($t(223) = 11.81$, $p < .001$). Thus, students were better able to monitor their science learning through the game, an essential skill for scientific literacy. Future research should continue to examine the benefits of GBLEs versus CBLEs in terms of benefits for motivation and learning.

Keywords: GBLE; STEM learning; Scientific Literacy; Metacognition; Reading Comprehension

1. Introduction

Well-designed GBLEs can benefit students' learning outcomes as well as their self-regulated learning (SRL) (Nietfeld, 2018). SRL encompasses cognitive, affective, and motivational elements (Clark et al., 2016; Connolly et al., 2012; Sailer & Homner, 2020). Characteristics such as goal-directed play and the ability to examine navigation paths and engaged strategy use suggest that GBLEs may be particularly well-suited for examining SRL elements. GBLEs provide a rich context to study and support students' SRL while augmenting typical traditional classroom environments. The goal of *Missions with Monty* is to support students' science literacy and comprehension monitoring, a key component of SRL. Students' display effective metacognition through comprehension monitoring, where students reflect about how well they understand the content that they are reading. Comprehension monitoring includes both evaluation and regulation activities when learning from text (Hacker, 1998). This includes a reader evaluating the extent to which they comprehend a text and subsequent attempts to regulate fix up strategies to improve any inadequacies with comprehension. Accurate comprehension monitoring is critical in order to judge what information needs to be reread or restudied and it is therefore important to train students to make accurate monitoring judgments (Huff & Nietfeld, 2009). *Missions with Monty* provides support for essential science content learning by equipping students with science literacy skills through comprehension monitoring strategy training. Educational games that emphasize science learning are prevalent; however, most do not attend to the importance of reading comprehension and comprehension monitoring.

Previous GBLEs have likely avoided supporting reading comprehension and monitoring given the many challenges that arise when teaching effective SRL strategies. Designing a game environment that can effectively scaffold the complex processes and strategies for SRL reading is exceptionally challenging as breakdowns in comprehension can occur for various reasons related to self-regulation, including deficits in prior knowledge

(McCarthy & McNamara, 2021), metacognitive monitoring (Follmer & Sperling, 2018), and motivational factors, such as low self-efficacy (Schraw et al., 2006), interest (Renninger & Hidi, 2015; Springer et al., 2017), or lack of a mastery-based approach (Seaton et al., 2014). SRL abilities develop in late elementary school and this provides a critical opportunity to intervene and support SRL with comprehension monitoring scaffolds (Tonks & Taboada, 2011). Regardless of the fact that little time is spent teaching effective comprehension monitoring strategies, students need to engage in successful reading and reading of science text (Blachowicz & Ogle, 2017). GLBEs are well-positioned to address this challenge; they offer a unique space to scaffold learning through explicit training, practice, and timely and customized feedback, all while providing the engaging elements that games offer.

Even though there is a deep and growing literature of metacognitive interventions and programs for upper elementary students in traditional classroom settings, there is surprisingly scant research focusing on metacognition in GBLEs (Braad, et al., 2020). Exceptions have included work in the *Crystal Island - Outbreak* GBLE where metacognitive response bias predicted game performance even after accounting for prior knowledge and motivational variables such as self-efficacy, interest, and goal orientation (Nietfeld, et al., 2014). Also in *Crystal Island*, Taub et al. (2018) used sequential pattern mining to distinguish between more and less efficient learners by revealing that more efficient learners tested fewer partially-relevant food items. Alternatively, Riemer and Schrader (2020) found that self-monitoring was related to learning outcomes only for students with a low or moderate competition preference when playing an entrepreneurship game. While these studies provide a useful start in the examination of metacognition in GBLEs, the field is open to the exploration of numerous topics, such as the role of metacognitive judgments in academic learning, that dominate much of the literature in traditional classroom environments.

2. Game Design

Missions with Monty (see Figure 1) was developed with a multidisciplinary team of educational psychologists, educational technology specialists, developers, curriculum specialists, and designers. The game was created to align with how teachers and students commonly use educational games: on laptops and tablets. Importantly, *Missions with Monty* was developed to support learners' science literacy skills and comprehension monitoring strategies within the context of learning from informational science texts. All science content in the game was developed to align with standards not only at the classroom and state level, but also nationally (e.g., Next Generation Science Standards). For example, informational texts cover content about ecosystems, energy, and Earth and human activity. The program's organization prioritizes science content presentation through texts embedded within 'missions' or units; across missions, a consistent narrative is present and endorsed by other game characters who are experts in their respective mission's content.

Figure 1: Reading text presented by animal researchers in *Missions with Monty*

Students start *Missions with Monty* by accessing an animated video that provides the game's narrative background. Students learn that they have been invited to work with Monty, a monitor lizard and famous scientist. Monty is also the founder of Wildlife University (WU), which is situated in a rainforest. WU was established not only as a base for Monty's research work, but also to support students' and professors' development of science literacy. The students and professors at WU are working hard to learn about ways to save their natural habitats. Once players arrive at WU, however, they quickly learn about two major problems: the students and professors at WU have become ill, and Monty is missing. Learners are tasked with two major gameplay goals: *1) learn as much as possible to figure out why the animals at WU are sick, and 2) identify who kidnapped Monty*. To accomplish these goals, players are informed that they will be supported by Monty's colleagues, who are experts in the science content students will learn.

Given the focus on SRL to support their learning and navigation, before the first mission, players complete a 'training camp' (see Figure 2) led by support-team 'professors' specializing in highlighting, summarization, and monitoring. Students were trained to selectively highlight main ideas to encourage a focus on quality, self-regulated highlighting, self-assessment, and revision (Leopold & Leutner, 2015). Summarization was also chosen to focus the player on succinctly recognizing and conveying the main idea of a text. Summarization has been shown to lead to improved comprehension monitoring accuracy (Thiede & Anderson, 2003). Of specific focus for the current study, monitoring accuracy assessed students' ability to calibrate their judgments of confidence with their actual performance on science items presented by game characters. Calibration as an index of metacognitive monitoring has been shown to relate strongly with academic performance and even predict efficiency in game-based learning contexts (Nietfeld, et al., 2006; Nietfeld, et al., 2014).

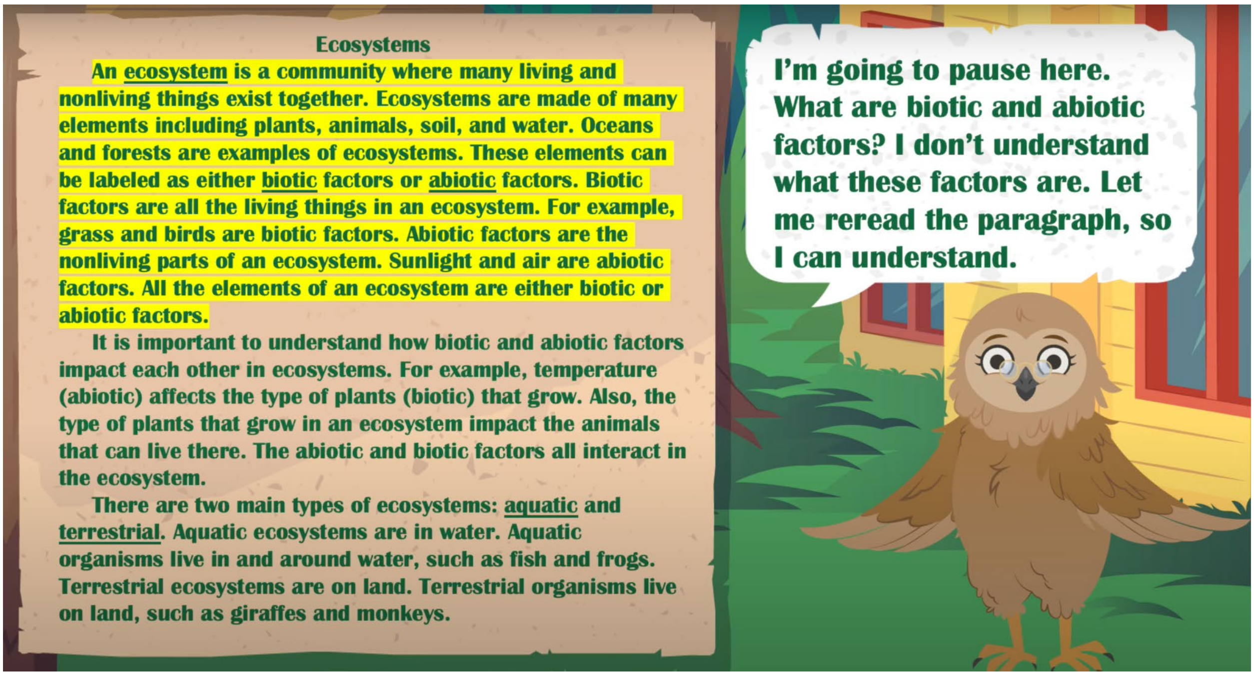

Figure 2: Monitoring training

After completing training camp, students embark on a total of three missions, which are framed as research camps in the rainforest, ocean, and savanna. Within each mission, students access various research sites that are further organized by content. At each research site, the assigned game character introduces learners to informational science texts and other gamified features in the respective environment. The texts and assessments employed in the gamified challenges were previously vetted in a validation study conducted within the grant project. In addition, the texts include additional sources of information, such as pictures, data charts, and representations. Students engage with the texts in a standard format at each research camp:

1. Read a text presented by a character and highlight the most important ideas.
2. Engage in a Highlighting Challenge where the player selectively highlights the most important sentences.
3. Engage in a Summary Challenge to identify the best summary from an expert-generated list.
4. Engage in a Knowledge Challenge to answer content items.
5. Engage in a Monitoring Challenge to provide confidence judgments.

6. Examine feedback on Knowledge Challenge answers and color-coded, customized feedback regarding their Monitoring Challenge judgments.
7. Complete the aforementioned steps for each passage in the research site.

In *Missions with Monty*, multiple-source understanding is essential to solve the overall problem of the game. Therefore, students engage in research site level and mission level multiple-source integration challenges called the Moment of Truth presented by the respective site and mission characters to demonstrate content mastery. They also participate in a final series of multiple-source questions presented by the Mission leader that function as the gateway to mastery of the Mission. An integrated understanding from multiple sources of information in the game is necessary in order to solve the problem of why animals are becoming sick (water contamination from fertilizer runoff – eutrophication).

All research sites and challenges are presented by characters within the game narrative. Responses to comprehension monitoring-based activities are assessed and translated into categorical game-based scores that translate to badges. Students receive immediate feedback on each task and have access to their progress in a Missions Journal. In the Missions Journal players track and then narrow their saved highlights, summaries, and images on a detective board where they can submit a hypothesis as to why animals are becoming ill. Students also select a suspect from their Suspect Board to try and guess who kidnapped Monty.

3. Research questions and Hypotheses

The current study targeted two primary research questions. First, *Do students learn science content from the game*? As a GBLE aligned with academic standards and used by educators to engage students to master those standards, it is essential that we establish that students can demonstrate learning on science outcomes. Second, *Does students' monitoring ability increase through the use of Missions with Monty, and are there differences in students' monitoring between those assigned Missions with Monty and those enrolled in a comparison condition?* Mayer (2015) presented three questions to guide GBLE research and in the current work we leveraged the media comparison question. Thus, a robust comparison condition was created for this study and administered to students through an engaging Google Forms format. The Google Forms were created to provide students with the same content included in *Missions with Monty* in an media-rich environment; however, students in the comparison condition did not receive any of the gamified features that students received in *Missions with Monty.* As such, students in the comparison condition accessed the Google Forms in the same standard format that students in *Missions with Monty* did, but they did not explore the content in a GBLE. For example, students in the comparison condition received static screenshots of the same characters from *Missions with Monty* producing dialogue in the Google Forms (See Figure 3), as opposed to dynamically interacting with them in the GBLE (e.g., navigating, clicking).

Figure 3: Screenshot of Google Form content

4. Methods

4.1 Participants

Fifth-grade students from four schools in North Carolina, USA participated in the study. Students had a mean age of 10.24 years (50.4% male, 46.0% female, 1.3% other, 2.2% preferred not to disclose). Demographics included students from 67.5% White, 11.2% Black or African American, 9.8 % Hispanic or Latin American, 0.9% Native American/Pacific Islander/Alaskan Native, 4.0% two or more races, and 1.8% Asian backgrounds along with 4.3% who preferred not to disclose their backgrounds. Intact classrooms were randomly assigned to either the GBLE or Google Form comparison; there were 144 students in the treatment condition and 80 students in the comparison condition.

4.2 Measures

Before gameplay participants completed a set of instruments including demographics, interest, and a science knowledge test derived from released test items from national and international tests related to the curriculum in the game. Students were also asked questions related to their digital game preferences and history. The targeted focus of this study, however, was the potential for the *Missions with Monty* to impact science learning and metacognitive monitoring ability. Science learning was measured through an 18-item multiple-choice assessment administered both pre and post intervention. A sample knowledge item included: *A lizard's tail breaks very easily. How does this help to protect the lizard? d) The lizard can leave pieces of its tail to mark its territory; b) During a food shortage, the lizard can break off its tail and eat it; c) The lizard can leave broken pieces of tail to fool predators; or d) The lizard's tail can break off if a predator attacks*. Participants rated their confidence at the item level on a 100-point scale (See Figure 4). Bias and accuracy scores were calculated as measures of metacognitive monitoring. Positive bias scores represented overconfidence, zero no bias, and negative scores represented underconfidence. Increased accuracy was represented by scores approaching zero, the number representing perfect judgment accuracy.

Figure 4: Confidence-rating interface in the GBLE

4.3 Procedures

The study was integrated within typical classroom instruction. Before engaging the content, students in both conditions completed the series of pre-test measures. Next, students in classrooms were randomly assigned to the GBLE or Google Form comparison condition. Since the study was integrated within typical classroom instruction, students completed *Missions with Monty* or the comparison condition at the pace determined by the teacher, typically taking place over about four weeks. After completing the programs, students in both conditions completed a series of post-test measures.

5. Results

Table 1 provides descriptive information for primary variables across both conditions. Analyses first established the two conditions did not differ significantly before intervention. No significant differences were found between groups on science knowledge at pretest ($t(222) = -.824$, $p = .41$, $d = 3.69$). Also, no significant differences were found between groups on average pre monitoring bias ($t(222) = 1.84$, $p = .07$, $d = .21$). Further, no significant differences were found between groups on average pre monitoring accuracy ($t(125.46) = .30$, $p = .77$, $d = .14$).

To answer our first question, *Do students learn science content from the (Missions with Monty) game*? A dependent *t*-test was conducted to examine pre and post knowledge scores in the treatment condition. Findings indicated that students' science knowledge significantly increased after gameplay, $t(143) = -13.21$, $p < .001$. Our second question asked, *Does students' monitoring ability increase through the use of Missions with Monty, and are there differences in students' monitoring between those assigned Missions with Monty and those enrolled in a comparison condition?* Students in the treatment condition became significantly less biased towards overconfidence $t(143) = 3.85$, $p < .001$, and significantly increased the accuracy of their judgments $t(143) = 11.52$, $p < .001$. A univariate analysis next tested the differences between conditions in monitoring bias when controlling for pre monitoring bias and found no significant differences ($F = 3.13$, $p = .08$). A univariate analysis next tested the differences between conditions in monitoring accuracy after controlling for pre-monitoring accuracy and found no significant differences ($F = 1.12$, $p = .29$). Importantly, significant gains for science learning were found across conditions, $t(223) = 13.67$, $p < .001$. Findings revealed that students in both conditions demonstrated judgments that were significantly less biased towards overconfidence $t(223) = 4.44$, $p < .001$ and accuracy $t(223) = 11.81$, $p < .001$.

Table 1: Means and standard deviations

	Mean	Standard Deviation
Science Pre Knowledge Comparison	10.38	4.07
Science Pre Knowledge Treatment	10.80	3.46
Pre Monitoring Bias Comparison	.15	.23
Pre Monitoring Bias Treatment	.09	.20
Pre Monitoring Accuracy Comparison	.38	.17
Pre Monitoring Accuracy Treatment	.37	.12
Science Post Knowledge Comparison	12.80	3.32
Science Post Knowledge Treatment	13.60	2.69
Post Monitoring Bias Comparison	0.08	.17
Post Monitoring Bias Treatment	.03	.15
Post Monitoring Accuracy Comparison	.23	.18
Post Monitoring Accuracy Treatment	.25	.13

6. Implications and Future Directions

The current study is contextualized within recent research attempting to develop both conceptual knowledge and SRL skills within GBLEs. Few GBLEs have been created to develop skills in multiple academic domains such as *Missions with Monty,* thus requiring increased demands on SRL skills and for pedagogical tools necessary for appropriate scaffolds and training. Findings demonstrated that students learned science content aligned with established academic standards through *Missions with Monty* gameplay. Students also gained in their ability to monitor their comprehension as demonstrated by both decreased overconfidence bias and increased accuracy. These findings endorse the benefits for GBLEs, demonstrate the further potential for learning standards-aligned academic content in GBLEs, and support the potential for GBLEs to facilitate effective SRL.

For teaching practice these findings provide evidence justifying the use of programs such as *Missions with Monty* to support both science learning and comprehension monitoring processes. *Missions with Monty* can be used in parallel with typical classroom instruction to augment strategy instruction, content learning, and individualized feedback. The current findings add to the rather sparse literature on promoting SRL through GBLEs.

Demonstrating significant improvement in monitoring accuracy in traditional classrooms has been very challenging and typically takes distributed practice with clear feedback over a number of weeks (Foster et al., 2017; Händel et al., 2020; Nietfeld, et al., 2006;). In the current study, we were able to demonstrate significant pre to post changes in monitoring accuracy with greater efficiency, a finding not yet reported in this literature. Future research should test the benefits of practice and varied feedback models on learning, comprehension monitoring, and transfer within STEM learning environments using GBLEs. Studies should also examine how to more fully integrate GBLEs into typical classroom instruction and to test immediate and customized feedback capabilities that are not easily provided without the use of such technologies.

References

Adams, D. M., Mayer, R. E., MacNamara, A., Koenig, A., & Wainess, R. (2012). Narrative games for learning: Testing the discovery and narrative hypotheses. *Journal of Educational Psychology, 104,* 235–249.

Blachowicz, C., & Ogle, D. (2017). *Reading comprehension: Strategies for independent learners*. New York, NY. Guilford Publications.

Braad, E., Degens, N., IJsselsteijn, W. A. (2020). Designing for metacognition in game-based learning: A qualitative review. *Translational Issues in Psychological Science, 6(1)*, 53-69.

Braten, I., Stromso, H. I., & Britt, M. A. (2009). Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly, 44*, 6-28.

Britt, M. A., & Rouet, J. F. (2012). Learning with multiple documents: Component skills and their acquisition. In J. R. Kirby, & M. J. Lawson (Eds.), *Enhancing the quality of learning: Dispositions, instruction, and learning processes* (pp. 276-314). New York: Cambridge University Press.

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research, 86*, 79-122.Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers and Education, 59*, 661-686.

Follmer, D. J., & Sperling, R. A. (2018). Interactions between reader and text: Contributions of cognitive processes, strategy use, and text cohesion to comprehension of expository science text. *Learning and Individual Differences*, *67*, 177-187.

Foster, N. L., Was, C. A., Dunlosky, J., & Isaacson, R. M. (2017). Even after thirteen class exams, students are still overconfident: The role of memory for past exam performance in student predictions. *Metacognition and Learning, 12*, 1-19.

Hacker, D. J. (1998). Self-regulated comprehension during normal reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 165-191). Mahwah, NJ: Erlbaum.

Händel, M., Harder, B., & Dresel, M. (2020). Enhanced monitoring accuracy and test performance: Incremental effects of judgment training over and above repeated testing. *Learning and Instruction, 65,* 101245. https://doi.org/10.1016/j.learninstruc.2019.101245

Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning, 4*, 161-176.

Hung, W., & van Eck, R. (2010). Aligning problem solving and gameplay: A model for future research and design. In R. van Eck (Ed.), *Interdisciplinary models and tools for serious games: Emerging concepts and future directions* (pp. 227–263). Hershey, PA: IGI Global.

Jackson, G. T., & McNamara, D. S. (2013). Motivation and performance in a game-based intelligent tutoring system. *Journal of Educational Psychology, 105*, 1036-1049.

Leopold, C., & Leutner, D. (2015). Improving students' science text comprehension through metacognitive self-regulation when applying learning strategies. *Metacognition and Learning, 10*, 313-346.

Mayer, R. E. (2015). On the need for research evidence to guide the design of computer games for learning. *Educational Psychologist, 50*(4), 349-353.

McCarthy, K. S., & McNamara, D. S. (2021). The multidimensional knowledge in text comprehension framework. *Educational Psychologist*, 1-19.

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

Nietfeld, J. L. (2018). The role of self-regulated learning in digital games. In D. Schunk & J. Greene (Eds.), *Handbook of Self-Regulation of Learning and Performance (2nd ed., pp. 271-284)*. Routledge.

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance and monitoring accuracy. *Metacognition and Learning, 2,* 159-179.

Nietfeld, J. L., Shores, L. R., & Hoffmann, K. F. (2014). Self-regulation and gender within a game-based learning environment, *Journal of Educational Psychology*, *106*, 961-973.

Renninger, K. A., & Hidi, S. E. (2015). *The power of interest for motivation and engagement*. Routledge.

Riemer, V., & Schrader, C. (2020): Playing to learn or to win? The role of students' competition preference on self-monitoring and learning outcome when learning with a serious game, *Interactive Learning Environments*, DOI: 10.1080/10494820.2020.1752741

Sailer, M., & Homner, L. (2020). The gamification of learning : A meta-analysis. *Educational Psychology Review, 32*, 77-112.

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review, 7*, 351-371.
Seaton, M., Parker, P., Marsh, H. W., Craven, R. G., & Yeung, A. S. (2014). The reciprocal relations between self-concept, motivation and achievement: juxtaposing academic self-concept and achievement goal orientations for mathematics success. *Educational psychology*, *34*, 49-72.
Springer, S. E., Dole, J. A., & Hacker, D. J. (2017). The role of interest in reading comprehension. In S. E. Israel (Ed.), *Handbook of research on reading comprehension* (pp. 519–542). The Guilford Press.
Taub, M., Azevedo, R., Bradbury, A. E., Millar, G. C., & Lester, J. (2018) Using sequence mining to reveal the efficiency in scientific reasoning during STEM learning with a game-based learning environment. *Learning and Instruction, 54*, 93-103.
Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95*, 66-73.
Tonks, S. M., & Taboada, A. (2011). Developing self-regulated readers through instruction for reading engagement. In B. J. Zimmerman, & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 173-186). New York: Routledge.