JOURNAL OF LEARNING ANALYTICS

SOLAR SOCIETY for LEARNING ANALYTICS RESEARCH

Volume 8(2), 51–72. https://doi.org/10.18608/jla.2021.7371

Game-Based Learning Analytics for Supporting Adolescents' Reflection

Elizabeth B. Cloude[1], Dan Carpenter[2], Daryn A. Dever[3], Roger Azevedo[4], James Lester[5]

Abstract

Reflection is critical for adolescents' problem solving and learning in game-based learning environments (GBLEs). Yet challenges exist in the literature because most studies lack a theoretical perspective and clear operational definition to inform how and when reflection should be scaffolded during game-based learning. In this paper, we address these issues by studying the quantity and quality of 120 adolescents' written reflections and their relation to their learning and problem solving with Crystal Island, a GBLE. Specifically, we (1) define reflection and how it relates to skill and knowledge acquisition; (2) review studies examining reflection and its relation to problem solving and learning with emerging technologies; and (3) provide direction for building reflection prompts into GBLEs that are aligned with the learning goals built into the learning session (e.g., learn about microbiology versus successfully solve a problem) to maximize adolescents' reflection, learning, and performance. Overall, our findings emphasize how important it is to examine not only the quantity of reflection but also the depth of written reflection as it relates to specific learning goals. We discuss the implications of using game-learning analytics to guide instructional decision making in the classroom.

Notes for Practice

- Reflection is essential for effective problem solving and learning in game-based learning environments (GBLEs).
- Multimodal data captured during game-based learning may provide insight into the quantity and quality of adolescents' reflections and their relation to learning and performance with GBLEs via game-learning analytics.
- Our findings suggest that designing reflection prompts based on learning goals in GBLEs provides insight into adolescents' learning and problem solving to guide instructional decision making in the classroom to support reflection, learning, and performance.

Keywords

reflection, game-learning analytics, adolescents, problem solving, knowledge acquisition

Submitted: 24/09/20 — **Accepted:** 30/03/21 — **Published:** 03/09/21

Corresponding author [1] *Email: elizabeth.cloude@knights.ucf.edu Address: Department of Learning Sciences and Educational Research, University of Central Florida, 4000 Central Florida Boulevard, Orlando, FL, 32816 USA. ORCID ID: https://orcid.org/0000-0002-7599-6768*

[2] *Email: dcarpen2@ncsu.edu Address: Department of Computer Science, North Carolina State University, 890 Oval Drive, Raleigh, NC, USA.*

[3] *Email: ddever@knights.ucf.edu Address: School of Modeling, Simulation, and Training, University of Central Florida, 3100 Technology Pkwy, Orlando, FL, USA.*

[4] *Email: roger.azevedo@ucf.edu Address: School of Modeling, Simulation, and Training, University of Central Florida, 3100 Technology Pkwy, Orlando, FL, USA.*

[5] *Email: lester@ncsu.edu Address: Department of Computer Science, North Carolina State University, 890 Oval Drive, Raleigh, NC, USA.*

1. Introduction

According to American philosopher John Dewey, meaningful experiences are not possible without some element of reflection (Dewey, 1923, 1933). Dewey (1933) suggests that reflection is the process of actively and carefully appraising situations or experiences that require learners to step back, observe, and meaningfully contemplate how they solve problems and whether a particular set of problem-solving strategies is effective for achieving their goals (Tarricone, 2011; Dewey, 1933). Thus, reflection is essential for effective problem solving because it promotes new knowledge and higher-order thinking skills (Dewey,

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

1933), which have significant bearing on learning and performance (Alzaid & Hsiao, 2018; Izu & Alexander, 2018; Luo & Baaki, 2019; Patel, Baker, & Scherer, 2019; Tarricone, 2011). Yet studies find that most learners in the United States do not have the skills for successful problem solving (Azevedo, Taub, & Mudrick, 2018; Azevedo, Feyzi-Behnagh, Duffy, Harley, & Trevors, 2012; NRC, 2012; NASEM, 2018; OECD, 2016), leading us to ask whether this insufficiency stems from a lack of reflecting. Since reflection is not taught in US curricula (NASEM, 2018), it is plausible that learners are not reflecting in order to cultivate skills that contribute to solving real-world problems (Moshman, 2013; Wang, Chen, Lin, & Hong, 2017). To deal with the educational challenges adolescents face (i.e., gaining the knowledge and skills needed to meet the demands of the 21st century), researchers have built game-based learning environments (GBLEs), such as Crystal Island, that incorporate features (e.g., narrative, pedagogical agents) designed to scaffold skill acquisition while simultaneously stimulating motivation and engagement (Taub, Sawyer, et al., 2020). Clark, Tanner-Smith, and Killingsworth (2016) found that, out of 57 studies, participants learning with GBLEs had consistently higher learning outcomes than those who learned in conventional settings without GBLEs ($\bar{g} = 0.33$, 95% CI [0.19, 0.48], $\tau^2 = 0.28$). Similar results were found by Mayer (2014) and Qian and Clark (2016). However, a study by Ak and Kutlu (2017) comparing traditional settings, 2-D GBLEs, and 3-D GBLEs found no differences in learning gains across the three learning environments ($ps > 0.05$). A possible explanation for the mixed findings could be that, while GBLE features are designed to foster higher-order thinking, most game features fail to scaffold reflection, which is important for acquiring knowledge and cultivating effective problem-solving skills in dynamic environments (Tarricone, 2011). This could be in part due to reflection often being vaguely defined and lacking a theoretical perspective, a dearth of studies that use multimodal learning analytics (Mangaroska, Sharma, Gasevic, & Giannakos, 2020; Geden et al., 2020), and the fact that game features (e.g., prompts) are not intentionally designed to scaffold reflection (Taub, Azevedo, Bradbury, & Mudrick, 2020). In this paper, gaps are addressed using a theoretical perspective and multimodal data on reflection to examine whether the quantity and quality of reflections were related to two learning goals presented within a GBLE, Crystal Island. Specifically, we (1) describe the challenges associated with how reflection has been previously defined and provide a clear definition that is supported by theoretical and empirical work; (2) highlight previous studies that have captured and analyzed reflection with emerging technologies; and finally (3) describe a model of reflection (McAlpine, Weston, Beauchamp, Wiseman, & Beauchamp, 1999) that emphasizes reflection as a function of achieving learning goals, offering a unique perspective on studying reflection with GBLEs, which implies that instructional decision making in the classroom should be augmented to scaffold adolescents' learning and problem solving.

1.1 What Is Reflection?

Significant challenges exist in the literature because reflection has often been defined interchangeably with introspection and metacognition (Brown, Bransford, Ferrara, & Campione, 1983; Flavell, 1979; Pintrich, 2002; Rosenthal, 2000). While reflection relies on introspection and serves as the foundation for metacognition, these terms are different (Tarricone, 2011). Specifically, reflection is the process of purposeful contemplation or focused thinking, whereas introspection is the process of looking within such that one draws awareness to one's thoughts, feelings, and reactions to stimuli (Tarricone, 2011). On the other hand, metacognition is the act of monitoring and controlling cognition (Flavell, 1979), while reflection is the act of observing, inspecting, and contemplating (1) a belief or supposed form of knowledge, (2) the evidence that supports said belief or knowledge, and (3) the conclusions one draws from observing and inspecting one's belief or knowledge.

Moreover, the literature refers to reflection as the building block of higher-order thinking (e.g., metacognitive monitoring, effective problem solving) (Tarricone, 2011). In other words, self-knowledge that results from introspecting and reflecting fuels metacognitive knowledge because it encompasses the beliefs learners hold about themselves and how they exist in the world (Pintrich, 2002). Traditional theoretical perspectives explain self-knowledge as a vital ingredient in developing metacognitive knowledge (e.g., model of metamemory by Flavell (1979)), leading to a number of empirical studies supporting the idea that metacognitive knowledge is essential for metacognitive processes such as feelings of knowing, and that reflection contributes to enhanced metacognition and learning outcomes (Bannert & Reimann, 2012). Further, the importance of developing intentional and conscious reflection is critical for developing higher-order thinking skills and plays a role in learning (Azevedo et al., 2018; Alzaid & Hsiao, 2018; Izu & Alexander, 2018; Luo & Baaki, 2019; Patel et al., 2019; Tarricone, 2011; Dewey, 1923).

Next, we discuss studies examining the role of reflection in skill/knowledge acquisition with emerging technologies.

1.2 Reflection with Emerging Technologies

Most studies capture reflection using prompts that are direct (i.e., the learner is prompted to reflect on specific aspects of their learning, e.g., "What is the most important information you learned about X?") or open (i.e., the learner is simply prompted to reflect in general, e.g., "How is your learning overall going so far?") (Vrugte et al., 2015), elicited using written statements or pedagogical agents with emerging technologies (Carpenter, Geden, Rowe, Azevedo, & Lester, 2020; Geden et al., 2020; Taub, Azevedo, et al., 2020). For instance, a study by Wu and Looi (2012) prompted adolescents (ages 13–15 years) to reflect using pedagogical agents with Betty's Brain, an agent-enabled, learning-by-teaching environment (Biswas, Segedy, & Bunchongchit,

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

2016). Participants were required to study concepts and diagrams related to elementary economics and then teach the agents based on their understanding of the learning materials. Participants were assigned to one of three groups: generic prompt, specific prompt, and no prompt. The generic prompts were designed to scaffold participants to examine their perspectives, beliefs, and experiences by reflecting on metacognitive strategies and beliefs used (e.g., "What do you think about teaching and who is it for?", p. 342), while the specific prompts were designed to scaffold participants to achieve the learning objectives by reflecting on task- and domain-specific skills (e.g., "Can you explain the concepts you just taught me?", p. 342). When participants indicated something incorrect about a concept during teaching to the agent, the agent was triggered to prompt the learner to reflect. Specifically, Wu and Looi (2012) examined relations between the quality of written reflections (i.e., did the written statement illustrate reactivity, contemplation, or elaboration?), immediate learning, and transfer knowledge using knowledge assessments across the three conditions. Results showed that participants assigned to either generic- or specific-prompt groups had higher immediate learning outcomes than the no-prompt group; however, there were no differences in immediate learning outcomes between the generic-prompt and specific-prompt groups. Their findings also suggested that adolescents in the generic-prompt group performed better on the transfer test than the specific-prompt and no-prompt groups (Wu & Looi, 2012). Additionally, their results showed that participants in the specific-prompt group demonstrated more reactivity in their written statements, while participants in the generic-prompt group demonstrated more contemplation in their written statements. However, significant challenges exist in this study, particularly in the timing of the reflection prompts via agents, which were initiated only when participants demonstrated incorrect patterns in relation to the learning material. For instance, should learners be prompted to reflect on their knowledge *only* when they illustrate misconceptions? At which points should adolescents be prompted to reflect while learning and teaching complex materials, and how does this prompting relate to achieving learning goals presented within the environment?

Many different GBLEs have been created to foster specific skill sets (e.g., problem solving, scientific reasoning) or conceptual understanding for various domains (e.g., computer science or microbiology) using clearly structured learning goals built into the environment. Yet few studies design game features to scaffold reflection using a theoretical perspective (Taub, Azevedo, et al., 2020), presenting significant challenges for our understanding of reflection during game-based learning and its role in learning and problem solving as it relates to achieving learning goals. Moreno and Mayer (2005) conducted a study where participants designed plants capable of surviving in different climates with a GBLE. Some participants were assigned to a condition where they received corrective or explanatory feedback on problem solving from a pedagogical agent, while others were assigned to a condition in which they were asked to reflect by explaining their problem-solving solutions (e.g., "Why do you think that particular type of root/stem/leaf will survive in this environment?") (Moreno & Mayer, 2005, p. 123). The results showed that participants who were asked to reflect on their solutions demonstrated better retention and transfer than those in the corrective or explanatory feedback condition. However, this effect was only present if participants reflected on a correct solution. While these findings demonstrate the promise of pedagogical agents scaffolding reflection and the positive effect of reflection on learning, we highlight that the study did not apply a theoretical model explaining or defining reflection, drawing us to question whether the prompts were eliciting reflection. This issue is especially emphasized when there is no analysis of the quality of the reflection responses as it relates to achieving learning goals. Additional gaps exist because there was no theoretical underpinning determining *when* reflection should be elicited during game-based learning. Similarly, Fiorella and Mayer (2012) used paper-based worksheets to scaffold reflection and examined whether reflection prompts impacted solving electric-circuit problems with GBLEs. The prompts were designed to draw learners' attention to specific information that was related to the problem solution (e.g., "If you take away a resistor in serial, the flow rate increases.") (Fiorella & Mayer, 2012, p. 177). They found that participants who were prompted to reflect performed better on the transfer test than those who were not (Fiorella & Mayer, 2012). However, we would like to raise similar issues emphasized in Moreno and Mayer (2005). If prompts were designed to direct learners' attention to specific information for finding the solution, learners could be metacognitively monitoring instead of reflecting. More issues exist because drawing learners' attention to specific information does not mean that the learner is practising reflection. Without a clear definition of reflection, it is difficult to make conclusions that might inform relations between reflection and learning with GBLEs, especially when reflection responses are not evaluated based on the quality of the responses as they relate to learning goals. A study by Johnson and Mayer (2010) compares the effects of two different types of prompts on performance—open and directed—against a no-prompt condition. Learners in the open-prompt condition were required to generate self-explanations after every action during game-based learning, while learners in the directed-prompt condition selected from a list of reasons after every action. Results showed that learners in the directed-prompt condition performed significantly better on the transfer assessment than those in the no-prompt condition ($d = 1.20$). They also found that learners in the open-prompt condition did not significantly differ from those in the no-prompt condition (Johnson & Mayer, 2010). It is surprising that no effect of open-ended reflection on learning was found since previous studies have found that open-ended written reflection responses reveal reflection depth that relates to higher learning outcomes (Wu & Looi, 2012; Carpenter et al., 2020; Ullmann, 2019). Additional challenges exist, since the simple act of selecting reasons for initiating an

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

action does not necessarily elicit reflection, drawing us to question whether reflection played a role in the differences between the conditions.

Conversely, Vrugte et al. (2015) found that different reflection-prompt types did not impact performance. They examined written reflections elicited via prompts during learning about math with a GBLE to assess its relation to reasoning and learning. Specifically, four conditions were built into the system: reflection-prompt condition, reflection prompt with procedural information condition, procedural information condition with no reflection prompts, and a control condition. Results showed reasoning skills improved across all conditions, suggesting reflection prompts or procedural information did not have an effect on learning (Vrugte et al., 2015). While the authors argued that the reflection prompts were too demanding to have an impact on learning, we would like to raise the issue that examining the quality of the reflection responses (e.g., depth at which one reflects) may provide more insight into the findings of this study, particularly into why there was no effect of reflection prompts on performance. For instance, can we ensure that the learners were reflecting solely based the prompt they received and the number of times they completed a written response, i.e., the quantity of written reflection? Carpenter et al. (2020) demonstrated a novel technique to gauge the quality of reflection by examining adolescents' written reflection responses using machine learning to automatically detect the quality at which learners reflected and assess its relation to learning with Crystal Island, a GBLE designed with two learning goals: learn about microbiology concepts and solve a mysterious illness on the island (Taub, Azevedo, et al., 2020). Reflection quality was defined based on the depth of the written response, with emphasis placed on whether the participant was observing, contemplating, and inspecting their thoughts, ideas, beliefs, and perspectives as they related to achieving a learning goal and whether they generated a clear hypothesis (i.e., beliefs about a particular direction for their problem solution or learning) to inform their next set of actions. Three prompt types were built into Crystal Island, requiring learners to reflect on different aspects of their learning and problem solving during game-based learning: (1) the important information they had learned, (2) their current approach to finding a solution, and (3) a different solution approach. These prompts were triggered when participants completed actions deemed critical for achieving the learning goals (see Table 1 for details). Results showed that the average reflection quality across the three reflection types was positively related to learning outcomes (Carpenter et al., 2020). This study highlights the fact that examining the quality with which adolescents reflect provides insight into the role of reflection on learning with GBLEs, but the authors failed to consider the other learning goal, i.e., solving the mystery to study how reflection quality impacted successful problem solving. For this very reason, we argue that a model of reflection that accounts for the goals or objectives built into the learning environment needs to be adopted in order to gain insight into whether learners were observing, contemplating, and inspecting their beliefs, actions, thoughts, etc., in relation to achieving their learning goals. As highlighted in the discussion above, gaps remain in the literature about whether different reflection-prompt types (e.g., open, directed) impact adolescents' ability to practise reflection and how the quality of their reflections relates to learning and problem solving as they align with particular learning goals and objectives (Carpenter et al., 2020; Geden et al., 2020; Vrugte et al., 2015; Johnson & Mayer, 2010; Fiorella & Mayer, 2012). To address these challenges, we guided our research using McAlpine et al. (1999)'s model of reflection relative to other models since it emphasizes that learning goals guide reflection, which impacts reasoning, decision making, monitoring, and knowledge acquisition.

1.3 Model of Reflection

The model of reflection (McAlpine et al., 1999) describes six main components: goals, knowledge, action, monitoring, decision making, and a corridor of tolerance, where reflection functions as a continuous and dynamic interaction between knowledge and action (see Figure 1). Specifically, the model explains that reflection is driven by learning goals (e.g., solve the mysterious illness versus learn as much as possible about microbiology with Crystal Island). Once goals are identified and established, the learner constructs plans to achieve set goals based on their knowledge (i.e., cognitive structures ranging from surface to deep level that are built from training and previous experiences), subsequently leading to actions that are continuously revised and evaluated via metacognitive monitoring that is based on the learner's corridor of tolerance. The corridor of tolerance is a mechanism that determines whether monitoring will result in decision making that leads to change, such as a change in strategy use (Shavelson & Stern, 1981). It is hypothesized that the corridor of tolerance determines whether change should be implemented based on the learner's threshold for what is acceptable, such that the cues the learner is monitoring will fall within acceptable or unacceptable boundaries based on how the learner defines progress toward achieving their goals. Depending on where this evaluation falls, a decision will be made about adjusting or continuing actions based on observing, contemplating, and inspecting plans, ideas, hypotheses, etc., that the learner has developed (McAlpine et al., 1999) as they relate to learning goals. Further, during inquiry, reflection serves to control inference making and reasoning while searching for information to inform a solution (Dewey, 1933; McAlpine et al., 1999). It is critical to emphasize that in the model, learning goals drive inquiry, reflection, inference making, and reasoning. We argue that McAlpine and colleagues' (1999) model of reflection is appropriate for studying reflection during game-based learning, since GBLEs are designed with clearly structured learning goals that learners must complete in order to be successful (Plass, Mayer, & Homer, 2020). Plass and colleagues (2020) explain that "... games for learning may be defined as games with specific learning goals" (p. 3); thus, the influence of

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

goals on problem-solving actions and learning information is evident. Significant gaps exist in the literature—this model has not been used in game-based learning or learning analytics research (Kovanović et al., 2018; Scheffel, Drachsler, Stoyanov, & Specht, 2014; Carpenter et al., 2020; Geden et al., 2020; Perez-Colado, Perez-Colado, Freire-Moran, Martinez-Ortiz, & Fernandez-Manjon, 2017; Taub, Azevedo, et al., 2020). We address this gap by examining the role of the quantity and quality of reflection as they relate to learning goals structured within Crystal Island to advance our understanding of reflection during game-based learning. Our study is outlined below.

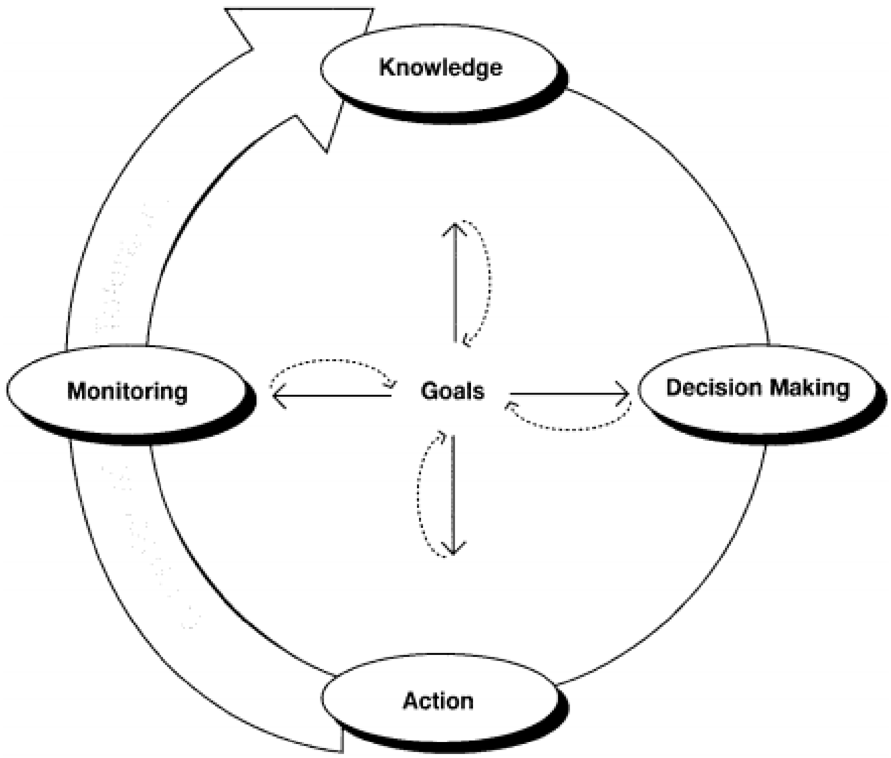

Figure 1. Model of reflection by McAlpine and colleagues (1999).

1.4 Current Study

In this study, we examined the role of reflection-prompt types on adolescents' problem-solving and learning outcomes with Crystal Island. Our research was guided by the model of reflection by McAlpine and colleagues (1999), because challenges that examine reflection prompts with GBLEs are prevalent in the literature (Carpenter et al., 2020; Moreno & Mayer, 2005; Fiorella & Mayer, 2012; Johnson & Mayer, 2010; Vrugte et al., 2015; Plass, Homer, & Kinzer, 2015; Taub, Azevedo, et al., 2020). Since studies show that reflection develops in adolescence (Moshman, 2011), it is critical to study how to effectively build reflection prompts within GBLEs to enhance learning and problem solving across learners. As such, the objectives of the current study are to analyze the effectiveness of three reflection-prompt types of scaffolding: (1) the important information previously learned; (2) the current approach to finding a solution; and (3) a different solution approach on two different learning goals, i.e., problem solving and knowledge acquisition with Crystal Island. To align our research questions with the model of reflection (McAlpine et al., 1999), we hypothesized that reflection-prompt types would impact learning and problem solving differently because they varied by goals (see Section 2.4, Coding and Scoring, for details).

- **What is the likelihood that a learner's ability to solve the problem is related to their quantity of reflections with Crystal Island?** While reflection positively impacted performance with GBLEs (Johnson & Mayer, 2010; Moreno & Mayer, 2005; Fiorella & Mayer, 2012), samples consist largely of young adults who may be more developmentally capable of reflecting with GBLEs. A study by Vrugte and colleagues (2015) found that reflection prompts did not affect adolescents' performance with GBLEs. Yet a study by Wu and Looi (2012) found a positive effect of adolescents' quantity of reflection on performance with a learning-by-teaching environment. Since the effect of reflection-prompt types built into Crystal Island on adolescents' problem solving has yet to be investigated, we adopted a non-directional hypothesis and expected the quantity of the three reflection-prompt types to impact adolescents' problem-solving performance.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

- **To what extent does the quantity of reflections predict post-test scores while controlling for pre-test scores with Crystal Island?** We expected the quantity of the three reflection-prompt types to impact adolescents' learning outcomes via pre-/post-test assessments based on the current literature (Johnson & Mayer, 2010; Moreno & Mayer, 2005; Fiorella & Mayer, 2012; Taub, Azevedo, et al., 2020); however, we adopted a non-directional hypothesis because of mixed findings on adolescents' reflection on learning and because the effect of the three reflection-prompt types built into Crystal Island on adolescents' learning outcomes has yet to be investigated.
- **What is the likelihood that a learner's ability to solve the problem is related to their average quality of reflections with Crystal Island?** Studies have found a positive effect of prompts on the depth to which adolescents reflect and acquire knowledge (Carpenter et al., 2020; Wu & Looi, 2012). Yet no studies have investigated the effect on adolescents' problem solving of reflection-prompt types that target different learning goals. As such, we adopted a non-directional hypothesis because the empirical evidence is unclear about adolescents' capacity to reflect with GBLEs, and gaps exist on the effectiveness of reflection-prompt types built into Crystal Island on adolescents' problem-solving performance. Specifically, we expected the average quality across reflection-prompt types to impact adolescents' problem-solving performance.
- **To what extent does the average quality of reflections predict post-test scores while controlling for pre-test scores with Crystal Island?** Studies have found a positive effect of reflection prompts on relations between adolescents' depth of reflection and learning (Carpenter et al., 2020; Wu & Looi, 2012). Yet no studies have investigated the effect on adolescents' learning of reflection-prompt types designed to target different learning goals. Because of this, we adopted a non-directional hypothesis and expected the average quality across the three reflection-prompt types to impact adolescents' learning outcomes.

2. Methods

2.1 Participants and Materials

A sample of 120 learners (51% female; age: $M = 13.57$, $SD = 0.54$) was recruited from several middle school classrooms to solve a mystery during their science class using Crystal Island, a narrative-based GBLE designed to foster (1) higher-order thinking skills and (2) microbiology knowledge (Rowe, Shores, Mott, & Lester, 2011; Taub, Sawyer, et al., 2020). Most participants identified as "White/Caucasian" (54%), while the remaining participants identified as "Black/African American" (27%), "Hispanic or Latino" (19%), or "Other" (16%). Most of the sample reported occasionally playing video games (33%), while others reported frequently (28%) and very frequently (20%) playing video games. The majority also reported possessing average video game skills (39%), while the remaining reported being skilled (29%) or of limited skills (14%). The majority of the sample also reported playing 0–2 hours of video games per week (38%), while others reported playing 5–10 (23%) and 3–5 hours (20%) of video games per week. None of the participants reported that they had learned with Crystal Island before the study.

A 21-item, four-option multiple-choice pre-/post-test assessment was administered before and after game-based learning with Crystal Island regardless of whether the mystery was solved to capture microbiology knowledge. The pre-/post-test assessments were created by the research team and middle school teachers, containing 12 factual (e.g., "What is the smallest type of living organism?") and 9 procedural (e.g., "What is the difference between bacterial and viral reproduction?") questions. Several self-report questionnaires were also administered to participants before and after game-based learning to gauge emotions, motivation, and cognitive load[1]. We obtained ethics committee approval before recruiting participants and collecting data. There was no experimental manipulation in this study because our aim was to explore how to best leverage Crystal Island to help learners engage in effective problem solving, knowledge construction, and higher-order thinking using reflection prompts and system features.

2.2 Crystal Island Learning Environment

The Crystal Island system is single-player GBLE built to create a rich, story-centred problem-solving experience using game features (e.g., tools) and embedded reflection prompts (Carpenter et al., 2020; Carpenter, Cloude, Rowe, Azevedo, & Lester, 2021; Geden et al., 2020). The science content provided in the game was aligned with the Standard Course of Study Essential Standards for Eighth-Grade Microbiology (McQuiggan, Goth, Ha, Rowe, & Lester, 2008). Upon starting the game, participants were required to adopt the role of a US Centers for Disease Control and Prevention agent who has been sent to a tropical island undergoing a recent illness outbreak. To successfully complete their mission, participants needed to identify (1) the pathogen, (2) the source of the pathogen (e.g., salmonellosis transmitted through eggs), and (3) a correct treatment plan. Tools and resources were designed into the system to provide clues and information about the illness, such as how the pathogen was

[1] We do not provide details on the self-report scales to maintain brevity since these data were not used in the analyses for this paper.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

contracted, using non-player characters (NPCs) like the camp nurse explaining patients' symptomatology; the ability to scan food items for pathogens; and the opportunity to read books, research articles, and posters that describe information about various pathogens and diseases (e.g., bacteria versus virus characteristics; Figure 2) and document hypotheses and findings using a diagnosis worksheet (Figure 3). The diagnosis worksheet was also a tool that participants used to submit their final diagnosis, source of contamination, and treatment solution.

The prompts embedded into the system were designed to foster reflection and were triggered using event-based production rules (Table 1). Further, the prompts asked participants to reflect on their problem solving once they completed actions deemed critical for solving the mystery via open-ended, short responses (e.g., "Please describe the most important things that you've learned so far, and what is your plan moving forward?") and rating scale items (i.e., "On a scale of 1–10, how well is your investigation going?"; Figure 4). Three types of reflection prompts were built into the system, corresponding to different learning goals: (1) progress plan, (2) solution approach, and (3) different solution approach, where each was designed to raise awareness of different aspects of problem solving. Specifically, the progress-plan prompts were triggered once a participant completed an action deemed critical for solving the illness, such as talking to the camp nurse to gather information on sick patients' symptoms. In response to actions like this, the learner was immediately prompted to describe the most important thing they had learned thus far in their problem solving toward identifying the pathogen. As such, the progress-plan prompt is distinct from the other prompts in that it was triggered during game-based learning, before participants indicated that they had reached a final solution, to foster their reflection on progress toward solving the mystery. Conversely, both solution-approach and different-solution-approach prompts were triggered when the learner submitted a correct diagnosis and treatment plan for the illness (i.e., solved the mystery), or if they ran out of class time. Thus, learners only completed one solution-approach prompt and one different-solution-approach prompt. Learners were first prompted to describe what problem-solving actions and information contributed to their success or failure (i.e., solution-approach). Afterwards, they were prompted to describe what problem-solving actions and strategies they would do differently (or the same) if given the chance to solve the mystery again (i.e., different-solution-approach). We would like to emphasize that the three reflection prompts were triggered at different points during game-based learning (i.e., before finding the solution and after) to scaffold reflection on and awareness of the participants' problem-solving approach and knowledge of microbiology.

2.3 Procedure

Before data collection, parents and participants provided written consent and assent to allow the latter's questionnaire, interaction log, and performance data to be used for research purposes. Consent and assent were obtained using forms distributed in person or electronically via email by researchers and middle school teachers. At least one day before playing Crystal Island, participants who provided both consent and assent completed a series of online questionnaires gauging experience with video games, emotions, metacognition, motivation, and microbiology knowledge via a 21-item multiple-choice pre-test assessment

Figure 2. Screenshots of GBLE, Crystal Island.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Table 1. Reflection triggers, prompts, and types during game-play

Triggers	Prompts	Types
After talking to the camp nurse for the first time	"Agent, it looks like you've spoken with the camp nurse. Before you continue, we'd like a report on your progress. In your own words, please describe the most important things that you've learned so far, and what is your plan moving forward?"	Progress plan
After viewing six virtual texts in the game	"Agent, it looks like you've found several materials that may be useful. Before you continue, we'd like a report on your progress. In your own words, please describe the most important things that you've learned so far, and what is your plan moving forward?"	Progress plan
After obtaining a positive test result in the virtual laboratory	"Agent, it looks like you found an object that tested positive for pathogenic contaminants. Before you continue, we'd like a report on your progress. In your own words, please describe the most important things that you've learned so far, and what is your plan moving forward?"	Progress plan
After submitting diagnosis worksheet to camp nurse and getting it wrong	"Agent, it looks like you are making progress on diagnosing the illness, but you're not quite there yet. In your own words, please describe the most important things that you've learned so far, and what is your plan moving forward?"	Progress plan
After solving the mystery	"Well done, Agent! You've saved everyone on the island. Now that you are finished, we would like to ask a couple of final questions. Please explain how you approached solving the mystery."	Solution approach
After solving the mystery	"If you were asked to solve a similar problem in the future, what would you do the same and/or differently?"	Different problem approach
After time expires, but the learner has not solved the mystery	"Thank you for playing Crystal Island. Now that you are finished, we would like to ask a couple of final questions. Please explain how you approached solving the mystery."	Solution approach
After time expires, but the learner has not solved the mystery	"If you were asked to solve a similar problem in the future, what would you do the same and/or differently?"	Different problem approach

using school computers during regular school hours. On the following day, participants were introduced to the game in their classroom by a researcher using a short video that explained the storyline of Crystal Island. Participants were told that the objective of the game was to solve a mysterious illness plaguing a remote, tropical island and learn about microbiology. Afterwards, participants interacted with the game on individual laptops provided by the research team for approximately 90 minutes during their normal science class time over the course of two class periods. Participants completed the study once the 90 minutes had passed on the second day, whether or not they solved the mystery illness. They did not receive the correct solution if they did not correctly solve the mystery to ensure that it did not confound their performance on the post-test knowledge assessment. Next, they completed a series of questionnaires similar to the pre-test session, including a 21-item multiple-choice post-test to capture microbiology knowledge.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Figure 3. Diagnosis worksheet.

Figure 4. Reflection prompt with written component on the left and scale on the right.

2.4 Coding and Scoring

2.4.1 Outcome Variables

McAlpine and colleagues' (1999) model of reflection guided our variable operational definitions. Since learning goals guide reflection according to this model, we created two outcome variables, which differed by learning goals: (1) microbiology knowledge acquisition and (2) successful problem solving. For instance, the goal of Crystal Island was to solve the mystery illness (i.e., problem solving), yet performing well on the post-test required learning as much as possible about microbiology, which may include information irrelevant to solving the mystery. Because of this, we defined outcome variables separately to study their relation to adolescents' reflection. To measure microbiology knowledge acquisition, i.e., learning, we used pre-/post-test scores that were defined based on the ratio of correct answers to total items on the assessment. Successful problem solving, i.e., performance, was defined based on whether the learner solved the mystery illness using a binary format (1 = solved, 0 = unsolved), which required the learner to identify not only the correct pathogen (salmonellosis) but also the source of contamination (transmitted via eggs) and to provide a correct treatment solution (rest) using the diagnosis worksheet.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

2.4.2 Reflection

To capture reflections, we extracted data from online behavioural traces, defining reflection in two ways: (1) quantity and (2) the average quality (or depth) of reflection responses. Overall, participants spent an average of 10 minutes ($SD = 3.08$) reflecting and 86.46 minutes ($SD = 10.60$) learning with Crystal Island. To define the quality (or depth), a rubric developed by Carpenter and colleagues (2020) was used to score the written reflection responses on a scale from one to five (1 = "no depth", 5 = "a reflective response with a high-quality sequence of abstract plans for problem solving"; see Table 2 for the rubric). While it is common for quantitative models of reflection to capture both depth (e.g., non-reflective, shallowly reflective, or highly reflective) and breadth (e.g., attending to feelings, validation, or justification) aspects of reflection (Kovanović et al., 2018; Ullmann, 2019), this rubric focused exclusively on depth. This was done because the reflections collected during game-based learning in Crystal Island were brief ($M = 20.22$ words, $SD = 15.56$) and therefore inherently limited in reflective breadth. The rubric of reflective depth was developed by two researchers using a grounded theory approach.

To determine what constituted a reflection rating of one, the two researchers searched through the reflections until several that seemed particularly weak were identified. Based on these reflections, the researchers determined that the weakest reflections were those that (1) lacked both a plan of action and commentary on relevant knowledge, (2) were too abstract to be very useful, (3) were largely unactionable, or (4) were entirely unrelated to the learning experience (e.g., "Yeah cool game I learned science"; see Table 2 for more examples). Next, the researchers selected several reflections that seemed exceptionally strong and used them as the basis for a reflection rating of five (e.g., "I will continue to test the foods the sick people touched or previously ate to see if it is contaminated"). The researchers determined that reflections that received a score of five either presented a clear hypothesis that was supported by strong evidence and reasoning or provided a high-quality abstraction of the problem that demonstrated an understanding of the most important information the student had obtained (see Table 2). A similar process was used to complete the rest of the rubric. In total, 20 reflections (four per reflection-depth rating) were used to develop the rubric (Carpenter et al., 2020). Once the rubric was developed, another 20 reflections were randomly sampled from the dataset and separately coded by each researcher. The ratings for these reflections were then discussed and any differences were reconciled, thus ensuring that both researchers had a shared understanding of the rubric. Finally, both researchers separately coded the remaining 708 reflections, and an intra-class correlation of 0.669 was achieved, indicating moderate inter-rater reliability (Cohen, 1960). The final reflection depth ratings used in this work were calculated by averaging the scores assigned by the two researchers ($M = 2.41$, $SD = 0.86$).

2.5 Statistical Analyses

The data were processed using a pipeline created in Python (Van Rossum & Drake, 2011) and then analyzed in R Version 3.6.2 (R Core Team, 2019). The packages `read xl` (Wickham & Bryan, 2019), `dplyr` (Wickham, François, Henry, & Müller, 2020), and `reshape2` (Wickham, 2007) were used for data manipulating and wrangling. The `fitdistrplus` and `logspline` packages were used to determine the distribution of the data before model building. To build different types of models, we used the `glm` and `lm` functions from the `stats` package (R Core Team, 2019). Next, the `stepAIC` function from the `MASS` package was used to conduct variable selection via the stepwise Akaike information criterion (AIC) (Venables & Ripley, 2002). The `base` and `lmtest` packages were used to access model indices, while the `ggplot2` package was used to visualize relationships among variables (Wickham, 2016).

For the analysis, binomial logistic regression models were calculated to assess relationships between the quantity and quality of reflection types and the likelihood that a learner solved the mystery. Multiple linear regression equations were used to assess relationships between the quantity and quality of reflection types and pre-/post-test proportional scores. The stepwise AIC was used to select the final models that demonstrated the lowest AIC for each research question. Since the stepwise AIC extends to generalized linear models, it was selected over other techniques (Yamashita, Yamashita, & Kamimura, 2007). Before the initial analyses, we removed 30 participants from the dataset because they did not complete the post-test assessment. We also investigated whether the quantity and quality of reflection types contained significant outliers using Grubbs's (1969) approach. Three participants were removed due to outlying data points, which were further emphasized using boxplots.

3. Results

Table 3 provides descriptive statistics. Alpha coefficients for pre-/post-test assessments met satisfactory reliability ($\alpha = 0.69$; (Cronbach, 1951)). Before model building, we conducted preliminary analyses to examine whether there were relationships between solving the mystery in Crystal Island and scores on pre-/post-test assessments. Two separate t-tests were calculated using a Bonferroni correction ($p < 0.05/2 = 0.025$) and revealed significant differences in pre-test scores ($t(115) = -2.33$, $p = 0.021$) and post-test scores ($t(112) = -2.65$, $p = 0.009$) between learners who solved (pre: $M = 0.35$; post: $M = 0.38$) or did not solve (pre: $M = 0.29$; post: $M = 0.30$; see Table 4 and Figure 5) the mystery. This suggests that an understanding of microbiology played a role in whether or not the learner was able to solve the mystery during learning with Crystal Island.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Table 2. Rubric used to annotate the depth of reflection-prompt responses (Carpenter et al., 2020)

Rating	Characteristics	Examples
1	Lacks both a plan and knowledge; abstract and largely meaningless; unactions.	"Each clue will help with solving the problem" or "Yeah cool game I learned science"
2	Presents a vague hypothesis or plan with no clear reasoning; simply restates information that was directly learned in the game without high-order thinking (e.g., inference making).	"That the illness causing the people being sick might be a pathogen" or "I found out that the egg has bacteria" or "I think I am going to talk to other people"
3	Presents a clear hypothesis or a plan, but does not provide reasoning behind it; demonstrates awareness of gaps in understanding and presents a plan to address those gaps; organizes the importance of their knowledge or problem-solving strategies.	"Getting more information about the food I think it has something to do with the food" or "The most important thing is how the illness is spreading"
4	Presents a clear hypothesis or plan with reasoning; provides reasoning of the situation with a plan; acknowledges what they have learned, why it is important, and what they plan to do with this information moving forward.	"I plan on questioning the cook as they know more about the food and how it could be contaminated with viruses or bacteria" or "I need to learn more about what the sick people do on a day-to-day schedule"
5	Presents both a clear hypothesis and plan with reasoning; presents a high-quality sequence of abstract plans for problem solving.	"I think that it might have to do with salmonella because when I tested the milk it was positive with pathogenic bacteria. I think that I will test things that can be contaminated" or "I will continue to test the foods the sick people touched or previously ate to see if it is contaminated"

Table 3. Descriptive statistics

Variables	*M*	*SD*	Range
Different problem approach quantity	1*	0	1
Different problem approach quality	1.97	1.04	4.5
Solution approach quantity	1*	0	1
Solution approach quality	2.65	1.42	5
Progress plan quantity	24*	23.04	36
Progress plan quality	2.21	0.61	4
Pre-test scores	0.32	0.13	0.62
Post-test scores	0.34	0.16	0.67

Note. $*$ = median.

3.1 Research Question 1: What Is the Likelihood That a Learner's Ability to Solve the Problem Is Related to Their Quantity of Reflections during Learning with Crystal Island?

AIC metrics indicated that the best model was present using one predictor: progress-plan reflection instances. A simple binomial logistic model was fit to the data (Table 5) to test the research hypothesis regarding the relationship between the likelihood that a

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Table 4. Frequencies on mystery solving

	Frequency	
	Yes	No
Solved the mystery	65 (54.62%)	55 (45.38%)

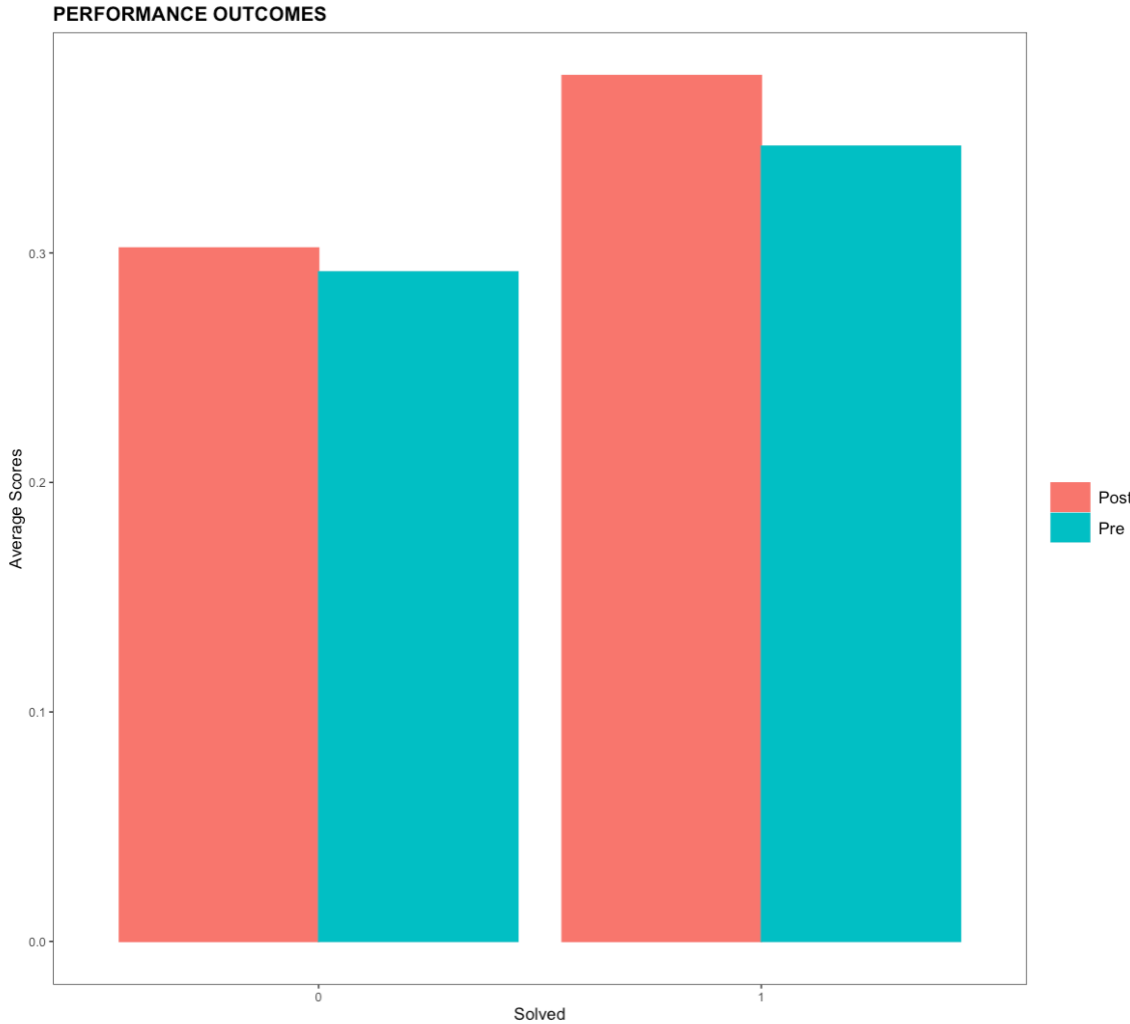

Figure 5. Differences in pre-/post-test scores between learners who solved and did not solve the mystery; 1 = solved, 0 = unsolved.

learner solves the mystery illness and the quantity of solution-approach and progress-plan reflections during learning with Crystal Island. We calculated the final model, which showed that Predicted logit of (Solved) $= -3.14 + (-0.06) * \text{Progress plans}$. According to the model, the likelihood of a learner solving the mystery illness was related to the quantity of completing progress-plan reflections ($p < 0.025$; Figure 5). In other words, the more progress-plan reflections were completed by the learner during game-based learning, the more likely the learner was to solve the mystery. Specifically, the odds of a learner solving the mystery illness were 0.94 times greater if they completed more progress-plan reflections than for a learner who completed fewer progress-plan reflections. This finding was partially consistent with our hypothesis and the model of reflection (McAlpine et al., 1999), where we expected the quantity of all reflection types to predict the likelihood that a learner solved the mystery. Please note that these findings are exploratory and have correlational implications. However, learners were not required to initiate specific actions during game-based learning that triggered reflection prompts, so the more often that learners initiated actions that were relevant to achieving the learning goals, the more often they were prompted to reflect.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Table 5. Logistic regression analysis of solving mystery by frequency of reflection types

Predictor	β	SE	p	e^{β}
Constant	−3.14	0.81	0.000	0.04
Progress plans	−0.06*	0.03	0.000	0.94
Test		χ^2	df	p
Overall model evaluation				
Likelihood ratio test		43.145	2	< 0.001
Goodness-of-fit test				
Hosmer & Lemeshow		8.5243	8	0.384
McFadden R^2	0.261			

Note. e^{β} = exponentiated beta or odds ratio; SE = standard error; $*p < 0.025$.

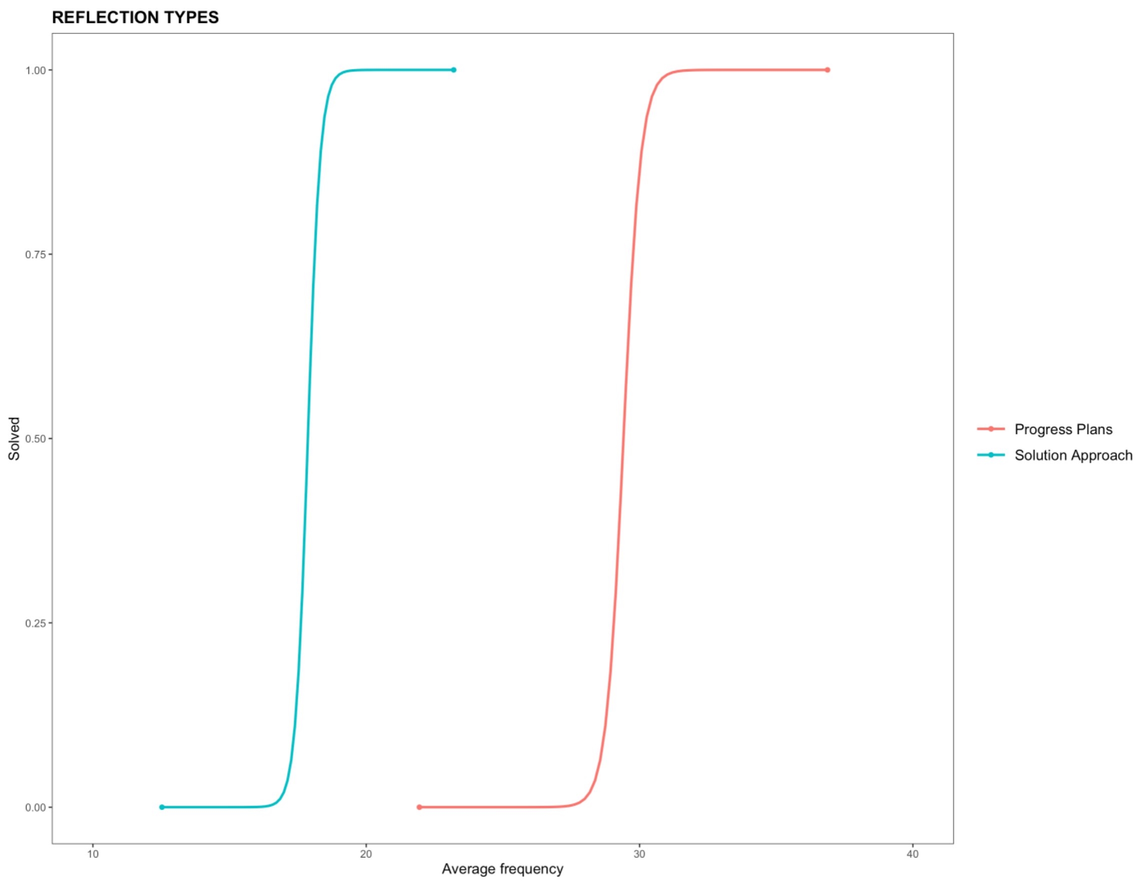

Figure 6. Line graph illustrating relationships between the average frequency of reflection types between learners who did and did not solve the mystery (solved = 1; unsolved = 0).

3.2 Research Question 2: To What Extent Does the Quantity of Reflections Predict Post-test Scores while Controlling for Pre-test Scores after Learning with Crystal Island?

Upon applying the AIC method, the best model indices were present using one predictor. A simple linear equation was fit to the data (Table 6) to test the research hypothesis regarding the relationships between pre- and post-test scores after learning with Crystal Island. The result showed that Predicted average of post-test scores $= 0.139 + (0.633) * \text{Pre-test scores}$. According to

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

the model, there were significant relationships between post-test scores and pre-test scores ($\beta = 0.63$, $p < 0.025$), but the model did not fit the data when including the quantity of reflection-type variables, suggesting that no relationship existed with post-test scores. Specifically, the fitted model estimated that the average post-test score increased by 0.633 for each point increase on the pre-test, where pre-test scores explained 27% of the variance in post-test scores (Figure 6). This finding was inconsistent with our hypothesis and previous research, where we expected the quantity of reflection types to predict post-test scores.

Table 6. Multiple linear regression analysis of post-test scores by frequency of reflection types and pre-test scores

Predictor	β	*SE*
Constant	1.39*	0.033
Pre-test scores	0.633*	0.095
Test		
F		44.44*
df		1,118
Adjusted R^2		0.267

Note. *SE* = standard error; *$p < 0.025$.

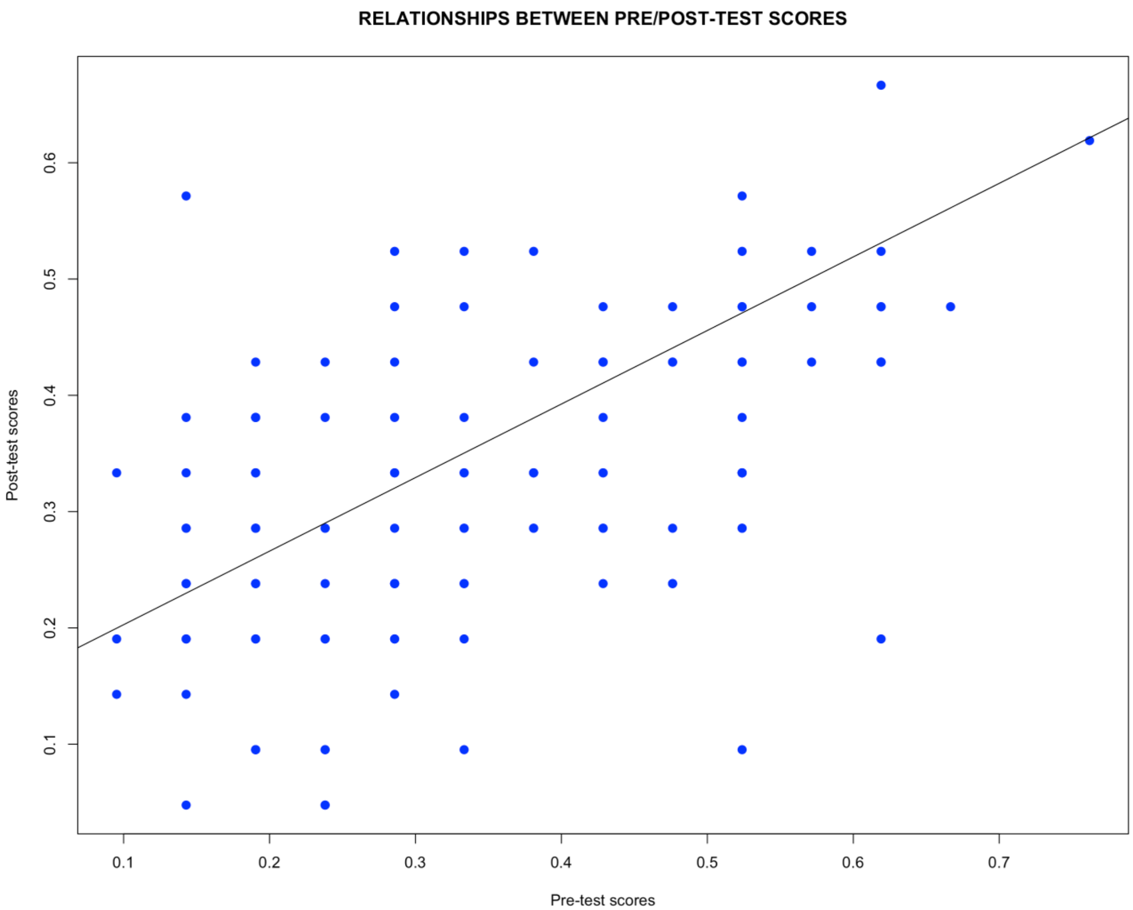

Figure 7. Scatterplot and regression line illustrating relations between pre- and post-test scores.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

3.3 Research Question 3: What Is the Likelihood That a Learner's Ability to Solve the Problem Is Related to Their Average Quality of Reflections during Learning with Crystal Island?

Upon applying the AIC method, the best model indices were present using one predictor. A one-predictor binomial logistic model was fit to the data (Table 7) to test the research hypothesis regarding the relationship between the likelihood that a learner solves the mystery illness and the average quality of solution-approach reflections during learning with Crystal Island. We calculated the final model, which showed that Predicted logit of (Solved) $= -2.20 + (0.90) * \text{Solution approach}$. According to the model, the likelihood of a learner solving the mystery illness was related to the average quality of their solution-approach reflections during game-based learning ($p < 0.025$; Figure 7). In other words, the higher the average quality of solution-approach reflections completed during game-based learning, the more likely a learner was to solve the mystery. Specifically, the odds of a learner solving the mystery illness were 2.46 times greater if they completed, on average, higher-quality solution-approach reflections than for a learner who completed, on average, lower-quality solution-approach reflections. This finding was partially consistent with our hypothesis and the model of reflection (McAlpine et al., 1999), where we expected the average quality of all reflection types to predict the likelihood that a learner solved the mystery.

Table 7. Logistic regression analysis of solving mystery by quality of reflection types

Predictor	β	SE	e^{β}
Constant	−2.20*	0.53	0.11
Solution approach	0.902*	0.19	2.46
Test		χ^2	p
Likelihood ratio test		43.145	< 0.001
Goodness-of-fit test			
Hosmer & Lemeshow		3.8462	0.8707
McFadden R^2	0.205		

Note. e^{β} = exponentiated beta or odds ratio; SE = standard error; $*p < 0.025$.

3.4 Research Question 4: To What Extent Does the Average Quality of Reflections Predict Post-test Scores while Controlling for Pre-test Scores after Learning with Crystal Island?

Upon applying the AIC method, the best model indices were present using two predictors. A multiple linear equation was fit to the data (Table 8) to test the research hypothesis regarding the relationships between the average quality of solution-approach and progress-plan reflections and post-test scores while controlling for pre-test scores after learning with Crystal Island. The results showed that Predicted average of post-test scores $= 0.012 + (0.533) * \text{Pre-test scores} + (0.054) * \text{Progress plan}$. According to the model, pre-test scores ($\beta = 0.53$, $p < 0.025$) and the average quality of progress-plan reflections ($\beta = 0.025$, $p < 0.025$) predicted post-test scores, but there was no relationship between the average quality of solution-approach reflections and post-test scores ($p > 0.025$). Specifically, the fitted model estimated that the average post-test score increased by 0.53 for each point increase on the pre-test and 0.05 for each point increase in the average quality of progress-plan reflections during game-based learning, where pre-test scores and average quality of progress-plan reflections explained 33% of the variance in post-test scores (Figures 8 and 9). This finding was partially consistent with our hypothesis and the model of reflection (McAlpine et al., 1999), where we expected the average quality of all reflection types to predict post-test scores.

4. Discussion

In this study, we examined the role of different reflection types on adolescents' problem-solving performance and knowledge acquisition with Crystal Island. Since challenges are prevalent in the literature examining reflection prompts with GBLEs, where a theoretical lens on and a clear operational definition of reflection are often lacking, our research was guided by the model of reflection by McAlpine and colleagues (1999) (Carpenter et al., 2020; Moreno & Mayer, 2005; Fiorella & Mayer, 2012; Johnson & Mayer, 2010; Vrugte et al., 2015; Plass et al., 2015; Taub, Azevedo, et al., 2020). The objective of the current study was to analyze data from adolescents to assess the effectiveness of three reflection-prompt types on different learning goals: (1) problem solving and (2) knowledge acquisition with Crystal Island. To align our research questions with the model

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

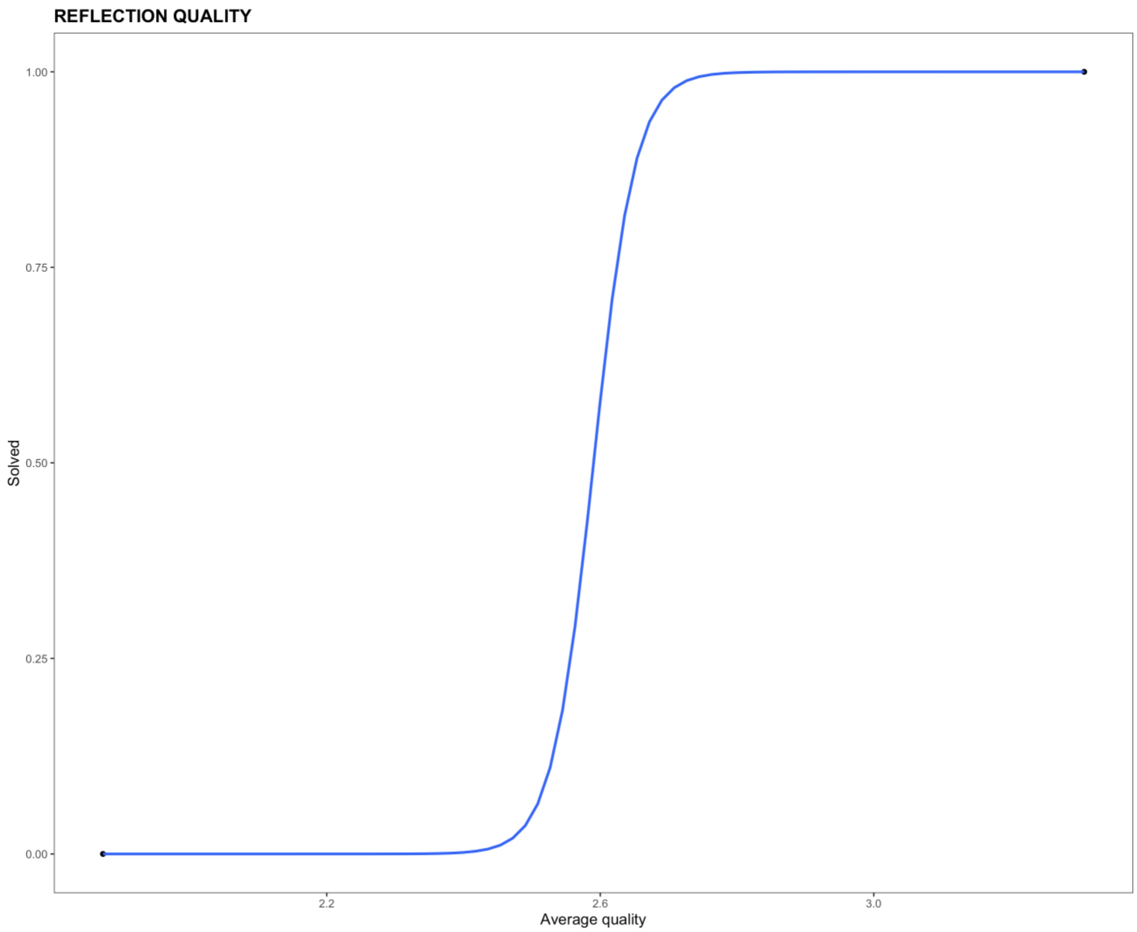

Figure 8. Line graph illustrating relationships between the average quality of solution-approach reflection between learners who did and did not solve the mystery (solved = 1; unsolved = 0).

Table 8. Multiple linear regression analysis of post-test scores by average quality of reflection types and pre-test scores

Predictor	β	*SE*
Constant	0.010	0.050
Pre-test scores	0.533*	0.095
Progress plans	0.054*	0.022
Solution approach	0.014	0.009
Test		
F		20.75*
df		3,116
Adjusted R^2		0.3324

Note. SE = standard error; $^*p < 0.025$.

of reflection (McAlpine et al., 1999), we hypothesized that reflection-prompt types impacted learning and problem solving differently because they varied by learning goals. It is important to note that research questions 1 and 3, and 2 and 4, were combined in this section to highlight findings by learning goal, i.e., solving the mystery versus learning about microbiology.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

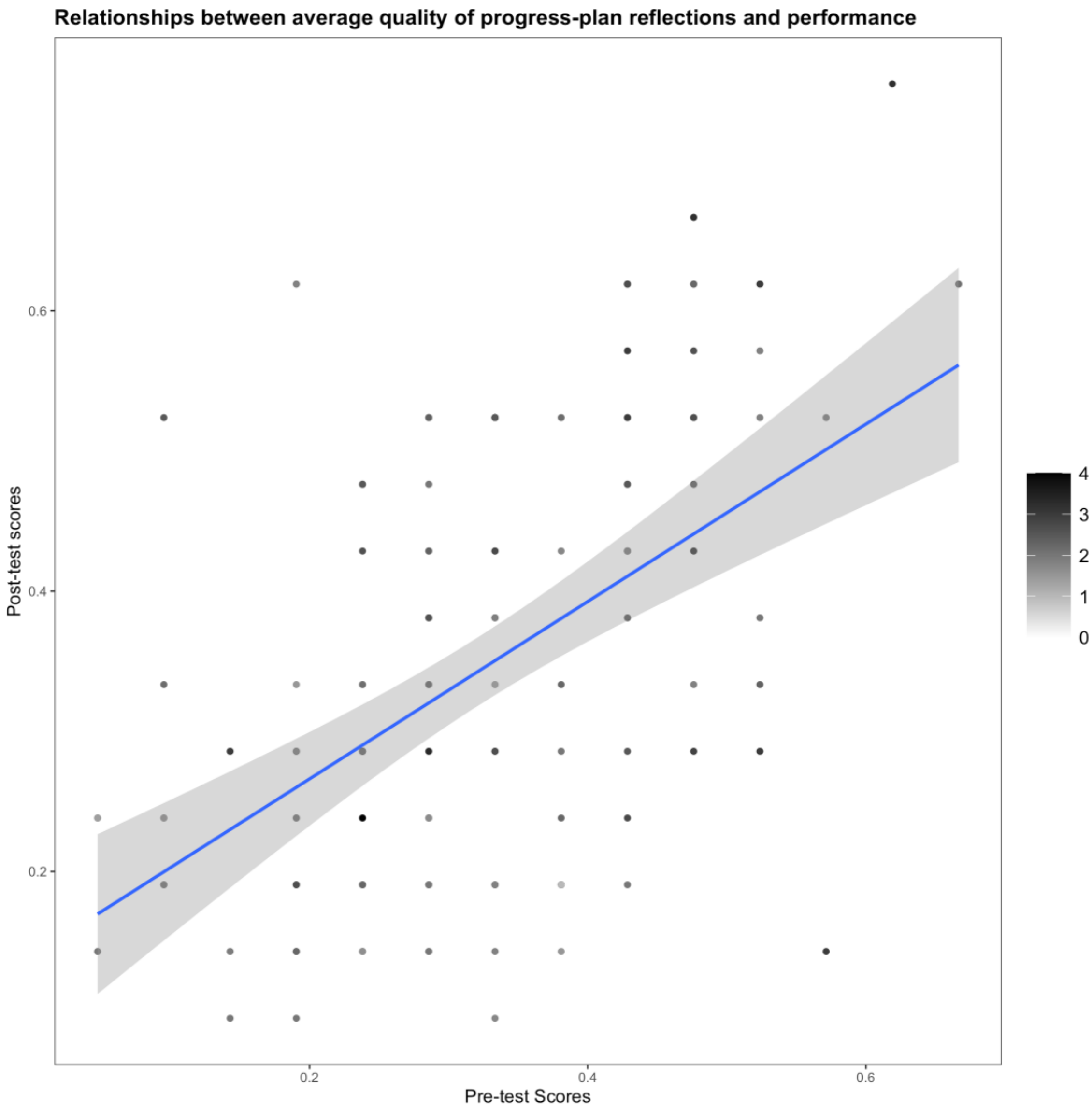

Figure 9. Scatterplot and regression line illustrating relations between pre- and post-test scores.

4.1 What Is the Likelihood That the Ability of a Learner to Solve the Mystery Illness Is Related to Both Their Quantity and Their Quality of Reflecting with Crystal Island?

The first research question examined whether the quantity of reflection-prompt types was related to the likelihood that a learner solved the mystery illness with Crystal Island. Our models found that the more often learners were prompted to reflect on their progress plans, the more likely they were to solve the problem in Crystal Island. These findings are partially consistent with our hypothesis, where we expected all reflection-prompt types to impact performance, since empirical evidence and the model of reflection (McAlpine et al., 1999) suggest that reflection enhances problem solving (Fiorella & Mayer, 2012; Johnson & Mayer, 2010; Moreno & Mayer, 2005; Wu & Looi, 2012). We would also like to note that these findings are exploratory and there was no control group to compare learners who were prompted to those who were not prompted at all. However, learners were not required to initiate specific actions during game-based learning, which triggered the reflection prompts, and so the more often learners initiated actions that were relevant to achieving the learning goals, the more often they were prompted to reflect. While we found that learners who reflected on their progress plans more often were more likely to solve the mystery, we did not find an effect of prompting learners to reflect on their solution approach or a different solution approach. A possible explanation for this could result from one of two things: (1) Prompting adolescents to reflect on their problem-solving approach after finishing the game might suggest that learners had active learning goals to solve the mystery; thus, according to McAlpine and colleagues (1999), there would be no drive to engage in reflection. (2) Prompting adolescents only one time might capture no variability

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

and thus leave little room for predictive models. As such, future studies should investigate whether designing GBLEs that prompt adolescents to contemplate their problem-solving approach at multiple instances during game-based learning affects reflection and problem-solving outcomes. If learners are actively pursuing a goal to solve the mystery, the prompts could be more impactful in scaffolding reflection and enhancing performance. Future studies should also consider the role that reflection plays in self-regulated learning by examining different theoretical perspectives such as that proposed by Zimmerman (2013) and capturing self-regulated learning processes and strategies over the course of game-based learning (Carpenter et al., 2021; Siadaty, Gasevic, & Hatala, 2016; Viberg, Khalil, & Baars, 2020)

The third research question examined whether the average quality of reflection-prompt types was related to the likelihood that a learner solved the illness with Crystal Island. Our models suggested that the higher the average quality (i.e., depth) of reflection responses on their solution approach, the more likely the learner was to solve the problem with Crystal Island. These findings highlight the importance of scaffolding adolescents' reflection on their approach to problem solving with GBLEs when they engage in problem-solving tasks. Further, the average quality of progress-plan and different-solution-approach prompts did not impact problem solving. These findings are partially consistent with our hypothesis, where we expected all reflection-prompt types to impact problem solving (Fiorella & Mayer, 2012; Johnson & Mayer, 2010; McAlpine et al., 1999; Moreno & Mayer, 2005; Wu & Looi, 2012). However, Vrugte et al. (2015) also found that some reflection prompts did not impact performance and suggested that the prompts were too demanding for adolescents. As such, a possible explanation for this finding could be that different-solution-approach prompts were too taxing for learners since they require knowledge of how to problem solve. Thus, future studies should aim to assess adolescents' problem-solving ability, which could explain why contemplating different problem-solving approaches was unrelated to performance.

In terms of progress plans, a possible explanation could be that the prompts were designed to scaffold the learner to contemplate the most important information they had learned about microbiology rather than about their problem solving. As such, could this prompt be related to the goal of learning about microbiology and, thus, according to the model of reflection (McAlpine et al., 1999), have discouraged the learner from reflecting on their progress toward solving the problem, potentially explaining why reflection quality didn't affect performance? A potential reason for the effect of the quantity of progress-plan prompts could be the number of times the GBLE was designed to prompt, which was considerably higher than for the other reflection-prompt types (Table 3). Further, could the quantity of progress-plan prompts simply have had an effect because they outnumbered the other prompts? Based on these findings, future studies should consider initiating solution-approach prompts at critical points during problem solving instead of progress-plan prompts to assess whether these prompts further enhance adolescents' reflection and problem solving. Future research should also aim to capture multimodal data during game-based learning because the information may provide more insight into the impact of prompts on adolescents' capacity to reflect. For example, how might facial expressions of emotions or eye-tracking data explain whether, when, and how reflection prompts are effective in scaffolding reflection and fostering successful problem solving? These findings have implications for the classroom, where the models could inform instruction and guide personalized interventions based on the quality and quantity of adolescents' reflection during game-based learning. However, considerable research is still required. Implementing game-learning analytics in the classroom for instructional decision making will also require additional support and technological resources, such as a dashboard to illustrate data visualizations for sense making that are currently unavailable to most teachers (Perez-Colado et al., 2017; Cloude, Dever, Wiedbusch, & Azevedo, 2020; Wiedbusch et al., 2021; Roll & Winne, 2015).

4.2 To What Extent Do the Quantity and Quality of Reflections Predict Post-test Scores while Controlling for Pre-test Scores with Crystal Island?

For our second research question, we examined whether the quantity of reflection-prompt types was related to post-test scores while controlling for pre-test scores. We found that the frequency of reflection prompts did not have an impact on pre-/post-test scores. Instead, prior knowledge was the best predictor of post-test scores. This finding was partially inconsistent with our hypothesis as well as previous research (Fiorella & Mayer, 2012; Johnson & Mayer, 2010; Moreno & Mayer, 2005; Wu & Looi, 2012), where we expected the quantity of reflection prompts to impact learning. A possible explanation could be that assessing the quantity of reflection has no bearing on whether learners engage in reflection with GBLEs. Future studies should aim to assess the depth of reflection and its relation to learning goals to study relations between reflection and outcomes with GBLEs. However, this finding highlights the importance of prior knowledge about the topic as it relates to post-test performance. For example, an instructor may want to pay close attention to learners who have demonstrated less understanding of the learning materials prior to engaging with a GBLE to inform their instructional decisions in the classroom, particularly in regard to scaffolding and intervening in a classroom with a lot of students, which is more often the rule than the exception. For instance, if an instructor knows that prior knowledge is related to learning outcomes, then they can monitor the subset of students who demonstrate less understanding.

For the final research question, we examined whether the average quality of reflection-prompt types was related to post-test scores while controlling for pre-test scores. We found that the average quality of progress-plan prompts and pre-test scores

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

positively predicted post-test scores. This finding was partially consistent with our hypothesis as well as previous research (Fiorella & Mayer, 2012; Johnson & Mayer, 2010; Moreno & Mayer, 2005; Wu & Looi, 2012), where we expected the quality of all reflection-prompt types to positively impact learning. A possible explanation for the lack of effect for solution-approach and different-solution-approach prompts could be that the pre-/post-test assessments were built to capture knowledge about microbiology, yet the objective of Crystal Island was to solve a problem in which solution-approach and different-solution-approach reflection-prompt types were created to scaffold (e.g., "Well done, Agent! You've saved everyone on the island. Now that you are finished, we would like to ask a couple of final questions. Please explain how you approached solving the mystery."). As such, the prompt types had no relevance to learning information about microbiology but rather scaffolded learners to reflect on problem solving. This explanation is aligned with the model of reflection (McAlpine et al., 1999), suggesting that learning goals play a role in driving reflection with GBLEs. Further, the progress-plan reflection prompts required the learner to think about the most important information acquired (e.g., "Agent, it looks like you are making progress on diagnosing the illness, but you're not quite there yet. In your own words, please describe the most important things that you've learned so far, and what is your plan moving forward?"), potentially targeting information covered on the pre-/post-test assessments. These findings provide directions for examining the role of reflection in performance and learning. When is the ideal time to prompt learners to reflect, and does that time vary based on the learning goal and environment? Researchers should also consider the role of adolescents' motivation, cognitive load, or level of knowledge of problem solving, since game-learning analytics data could help pinpoint when learners are not engaging in reflection (e.g., if the quality of solution-based reflection and motivation is low) to inform instructional decision making (Winne et al., 2019; Winne, 2017; Cloude et al., 2020; Wiedbusch et al., 2021).

4.3 Limitations

For this study, we did not assess individual characteristics outside of prior knowledge, such as motivation or interest, which may have played a role in adolescents' quantity and quality of reflection during game-based learning. Additionally, we did not capture the temporal components of learners' reflection, such as the amount of time engaging in reflection across the different prompt types, to assess its relation to performance and learning. Another confounding factor could have been adolescents' experience and familiarity with playing video games.

5. Concluding Statements

In this study, we examined the role of different reflection types on adolescents' problem-solving performance and knowledge acquisition with Crystal Island. Findings suggested that the quantity and quality of adolescents' reflection were related to problem solving and learning with a GBLE, but the effectiveness of different reflection-prompt types varied based on the learning goal they targeted. We emphasize the importance of adopting a theoretical perspective when studying reflection and highlight how learning goals built into GBLEs drive adolescents' reflection. Future studies should investigate the role of time and individual factors in reflection during game-based learning. Implications for building reflection prompts within GBLEs to enhance adolescents' reflection, problem solving, and learning, and using game-learning analytics to inform instruction in the classroom, are provided.

Declaration of Conflicting Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The publication of this article received financial support from the US National Science Foundation (DRL-1661202). Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of NSF.

References

Ak, O., & Kutlu, B. (2017). Comparing 2D and 3D game-based learning environments in terms of learning gains and student perceptions. *British Journal of Educational Technology*, *48*(1), 129–144. https://doi.org/10.1111/bjet.12346

Alzaid, M., & Hsiao, I.-H. (2018). Effectiveness of reflection on programming problem solving self-assessments. In *Proceedings of the ieee frontiers in education conference* (pp. 1–5). Piscataway, NJ, USA: IEEE. https://doi.org/10.1109/FIE.2018.8659245

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 & S. Land (Eds.), *Theoretical Foundations of Learning Environments* (pp. 171–197). New York, NY, USA: Routledge.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Azevedo, R., Taub, M., & Mudrick, N. (2018). Understanding and reasoning about real-time cognitive, affective, and metacognitive processes to foster self-regulation with advanced learning technologies. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of Self-Regulation of Learning and Performance* (pp. 254–270). New York, NY, USA: Routledge.

Bannert, M., & Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science*, *40*(1), 193–211. https://doi.org/10.1007/s11251-011-9167-4

Biswas, G., Segedy, J. R., & Bunchongchit, K. (2016). From design to implementation to practice a learning by teaching system: Betty's Brain. *International Journal of Artificial Intelligence in Education*, *26*(1), 350–364. https://doi.org/10.1007/s40593-015-0057-9

Brown, A., Bransford, J., Ferrara, R., & Campione, J. (1983). *Learning, Remembering, and Understanding* (Tech. Rep. No. 244). New York, NY, USA: Bolt, Beranek and Newman, Inc. and Center for the Study of Reading. Retrieved from https://files.eric.ed.gov/fulltext/ED217401.pdf

Carpenter, D., Cloude, E. B., Rowe, J., Azevedo, R., & Lester, J. (2021). Investigating student reflection during game-based learning in middle grades science. In *Proceedings of the 11th International Conference on Learning Analytics and Knowledge* (LAK 2021), 12–16 April 2021, online (pp. 280–291). New York, NY, USA: ACM.

Carpenter, D., Geden, M., Rowe, J., Azevedo, R., & Lester, J. (2020). Automated analysis of middle school students' written reflections during game-based learning. In I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Proceedings of the International Conference on Artificial Intelligence in Education* (AIED 2020), 6–10 July 2020, Utrecht, Netherlands (pp. 67–78). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-52237-7_6

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*(1), 79–122. https://doi.org/10.3102/0034654315582065

Cloude, E. B., Dever, D. A., Wiedbusch, M. D., & Azevedo, R. (2020). Quantifying scientific thinking using multichannel data with Crystal Island: Implications for individualized game-learning analytics. *Frontiers in Education*, *5*, 572546. https://doi.org/10.3389/feduc.2020.572546

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. Retrieved from https://doi.org/10.1177/001316446002000104

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Dewey, J. (1923). *Democracy and Education: An Introduction to the Philosophy of Education*. Macmillan.

Dewey, J. (1933). *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*. Lexington, MA, USA: D. C. Heath and Co.

Fiorella, L., & Mayer, R. E. (2012). Paper-based aids for learning with a computer-based game. *Journal of Educational Psychology*, *104*(4), 1074. https://doi.org/10.1037/a0028088

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Geden, M., Emerson, A., Carpenter, D., Rowe, J., Azevedo, R., & Lester, J. (2020). Predictive student modeling in game-based learning environments with word embedding representations of reflection. *International Journal of Artificial Intelligence in Education*, *7*(3), 1–23. https://doi.org/10.1007/s40593-020-00220-4

Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. *Technometrics*, *11*(1), 1–21. https://doi.org/10.1080/00401706.1969.10490657

Izu, C., & Alexander, B. (2018). Using unstructured practice plus reflection to develop programming/problem-solving fluency. In *Proceedings of the 20th Australasian Computing Education Conference* (ACE 2018), 30 January–2 February 2018, Brisbane, Australia (pp. 25–34). New York, NY, USA: ACM. https://doi.org/10.1145/3160489.3160496

Johnson, C. I., & Mayer, R. E. (2010). Applying the self-explanation principle to multimedia learning in a computer-based game-like environment. *Computers in Human Behavior*, *26*(6), 1246–1252. https://doi.org/10.1016/j.chb.2010.03.025

Kovanović, V., Joksimović, S., Mirriahi, N., Blaine, E., Gašević, D., Siemens, G., & Dawson, S. (2018). Understand students' self-reflections through learning analytics. In *Proceedings of the Eighth International Conference on Learning Analytics and Knowledge* (LAK 2018), 7–9 March 2018, Sydney, Australia (pp. 389–398). New York, NY, USA: ACM. https://doi.org/10.1145/3170358.3170374

Luo, T., & Baaki, J. (2019). Scaffolding problem-solving and instructional design processes: Engaging students in reflection-in-action and external representations in three online courses. In M. Boboc & S. Koç (Eds.), *Student-Centered Virtual Learning Environments in Higher Education* (pp. 40–69). Hershey, PA, USA: IGI Global. https://doi.org/10.4018/978-1-5225-5769-2.ch003

Mangaroska, K., Sharma, K., Gasevic, D., & Giannakos, M. (2020). Multimodal learning analytics to inform learning design: Lessons learned from computing education. *Journal of Learning Analytics*, *7*(3), 31–60. https://doi.org/10.18608/jla.2020.73.7

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Mayer, R. E. (2014). *Computer games for learning: An evidence-based approach*. Cambridge, MA, USA: MIT Press.

McAlpine, L., Weston, C., Beauchamp, C., Wiseman, C., & Beauchamp, J. (1999). Building a metacognitive model of reflection. *Higher Education*, *37*(2), 105–131. https://doi.org/10.1023/A:1003548425626

McQuiggan, S. W., Goth, J., Ha, E., Rowe, J. P., & Lester, J. C. (2008). Student note-taking in narrative-centered learning environments: Individual differences and learning effects. In *Proceedings of the International Conference on Intelligent Tutoring Systems* (ITS 2008), 23–27 June 2008, Montreal, QC, Canada (pp. 510–519). New York, NY, USA: Springer. Retrieved from https://doi.org/10.1007/978-3-540-69132-7$_5$4

Moreno, R., & Mayer, R. E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology*, *97*(1), 117–128. Retrieved from https://doi.org/10.1037/0022-0663.97.1.117

Moshman, D. (2011). *Adolescent Rationality and Development: Cognition, Morality, and Identity*. New York, NY, USA: Taylor & Francis.

Moshman, D. (2013). Adolescent rationality. In R. M. Lerner & J. B. Benson (Eds.), *Embodiment and Epigenesis: Theoretical and Methodological Issues in Understanding the Role of Biology within the Relational Developmental System* (Vol. 45, pp. 155–183). Amsterdam, Netherlands: Elsevier. Retrieved from https://doi.org/10.1016/B978-0-12-397946-9.00007-5

NASEM. (2018). *How People Learn II: Learners, Contexts, and Cultures*. Washington, DC, USA: The National Academies Press. Retrieved from https://www.doi.org/10.17226/24783

NRC. (2012). *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. Washington, DC, USA: The National Academies Press. https://doi.org/10.17226/13398

OECD. (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. Paris, France: OECD Publishing. Retrieved from https://www.doi.org/10.1787/9789264266490-en

Patel, N., Baker, S. G., & Scherer, L. D. (2019). Evaluating the cognitive reflection test as a measure of intuition/reflection, numeracy, and insight problem solving, and the implications for understanding real-world judgments and beliefs. *Journal of Experimental Psychology: General*, *148*(12), 2129–2153. https://doi.org/10.1037/xge0000592

Perez-Colado, I. J., Perez-Colado, V. M., Freire-Moran, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2017). Integrating learning analytics into a game authoring tool. In *Proceedings of the International Conference on Web-Based Learning* (ICWL 2017), 20–22 September 2017, Cape Town, South Africa (pp. 51–61). New York, NY, USA: Springer. https://doi.org/10.1007/978-3-319-66733-1_6

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory Into Practice*, *41*(4), 219–225. https://doi.org/10.1207/s15430421tip4104_3

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, *50*(4), 258–283. https://doi.org/10.1080/00461520.2015.1122533

Plass, J. L., Mayer, R. E., & Homer, B. D. (2020). *Handbook of Game-Based Learning*. Cambridge, MA, USA: MIT Press.

Qian, M., & Clark, K. R. (2016). Game-based learning and 21st century skills: A review of recent research. *Computers in Human Behavior*, *63*, 50–58. https://doi.org/10.1016/j.chb.2016.05.023

R Core Team. (2019). *R: A Language and Environment for Statistical Computing* [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, *2*(1), 7–12. Retrieved from https://doi.org/10.18608/jla.2015.21.2

Rosenthal, D. (2000). Introspection and self-interpretation. *Philosophical Topics*, *28*(2), 201–233. Retrieved from https://www.jstor.org/stable/43154687

Rowe, J. P., Shores, L. R., Mott, B. W., & Lester, J. C. (2011). Integrating learning, problem solving, and engagement in narrative-centered learning environments. *International Journal of Artificial Intelligence in Education*, *21*(1-2), 115–133. https://doi.org/10.3233/JAI-2011-019

Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality indicators for learning analytics. *Journal of Educational Technology & Society*, *17*(4), 117–132. Retrieved from https://www.jstor.org/stable/jeductechsoci.17.4.117

Shavelson, R. J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. *Review of Educational Research*, *51*(4), 455–498. https://doi.org/10.3102/00346543051004455

Siadaty, M., Gasevic, D., & Hatala, M. (2016). Trace-based micro-analytic measurement of self-regulated learning processes. *Journal of Learning Analytics*, *3*(1), 183–214. Retrieved from https://doi.org/10.18608/jla.2016.31.11

Tarricone, P. (2011). *The Taxonomy of Metacognition*. London, England, UK: Psychology Press. Retrieved from https://doi.org/10.4324/9780203830529

Taub, M., Azevedo, R., Bradbury, A. E., & Mudrick, N. V. (2020). Self-regulation and reflection during game-based learning. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of Game-Based Learning* (pp. 239–262). Cambridge, MA, USA: MIT Press.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Taub, M., Sawyer, R., Smith, A., Rowe, J., Azevedo, R., & Lester, J. (2020). The agency effect: The impact of student agency on learning, emotions, and problem-solving behaviors in a game-based learning environment. *Computers & Education*, *147*, 103781. https://doi.org/10.1016/j.compedu.2019.103781

Ullmann, T. D. (2019). Automated analysis of reflection in writing: Validating machine learning approaches. *International Journal of Artificial Intelligence in Education*, *29*(2), 217–257. https://doi.org/10.1007/s40593-019-00174-2

Van Rossum, G., & Drake, F. L. (2011). The Python Language Reference Manual [Computer software manual]. Network Theory Ltd.

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth ed.). New York, NY, USA: Springer. Retrieved from https://www.doi.org/10.1007/978-0-387-21706-2

Viberg, O., Khalil, M., & Baars, M. (2020). Self-regulated learning and learning analytics in online learning environments: A review of empirical research. In *Proceedings of the 10th international conference on learning analytics and knowledge* (pp. 524–533). New York, NY, USA: ACM. Retrieved from https://doi.org/10.1145/3375462.3375483

Vrugte, J. t., Jong, T. d., Wouters, P., Vandercruysse, S., Elen, J., & van Oostendorp, H. (2015). When a game supports prevocational math education but integrated reflection does not. *Journal of Computer Assisted Learning*, *31*(5), 462–480. https://doi.org/10.1111/jcal.12104

Wang, H.-H., Chen, H.-T., Lin, H.-S., & Hong, Z.-R. (2017). The effects of college students' positive thinking, learning motivation and self-regulation through a self-reflection intervention in Taiwan. *Higher Education, Research, & Development*, *36*(1), 201–216. https://doi.org/10.1080/07294360.2016.1176999

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, *21*(12), 1–20. Retrieved from http://www.jstatsoft.org/v21/i12/

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY, USA: Springer. https://doi.org/10.1007/978-3-319-24277-4

Wickham, H., & Bryan, J. (2019). *readxl: Read Excel Files* [Computer software manual]. (R package version 1.3.1) Retrieved from https://CRAN.R-project.org/package=readxl

Wickham, H., François, R., Henry, L., & Müller, K. (2020). *dplyr: A Grammar of Data Manipulation* [Computer software manual]. (R package version 0.8.5) https://doi.org/10.18637/jss.v021.i12

Wiedbusch, M. D., Kite, V., Yang, X., Park, S., Chi, M., Taub, M., & Azevedo, R. (2021). A theoretical and evidence-based conceptual design of MetaDash: An intelligent teacher dashboard to support teachers' decision making and students' self-regulated learning. *Frontiers in Education*, *6*, 14. https://doi.org/10.3389/feduc.2021.570229

Winne, P. H. (2017). Learning analytics for self-regulated learning. In C. Lang, G. Siemens, A. Wise, & D. Gašević (Eds.), *Handbook of Learning Analytics* (pp. 241–249). New York, NY, USA: Society for Learning Analytics Research. https://doi.org/10.18608/hla17.021

Winne, P. H., Teng, K., Chang, D., Lin, M. P.-C., Marzouk, Z., Nesbit, J. C., . . . Vytasek, J. (2019). nStudy: Software for learning analytics about processes for self-regulated learning. *Journal of Learning Analytics*, *6*(2), 95–106. Retrieved from https://doi.org/10.18608/jla.2019.62.7

Wu, L., & Looi, C.-K. (2012). Agent prompts: Scaffolding for productive reflection in an intelligent learning environment. *Journal of Educational Technology & Society*, *15*(1), 339–353. Retrieved from https://www.jstor.org/stable/jeductechsoci.15.1.339

Yamashita, T., Yamashita, K., & Kamimura, R. (2007). A stepwise AIC method for variable selection in linear regression. *Communications in Statistics—Theory and Methods*, *36*(13), 2395–2403. https://doi.org/10.1080/03610920701215639

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, *48*(3), 135–147. Retrieved from https://doi.org/10.1080/00461520.2013.794676

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)