Integrating Learning, Problem Solving, and Engagement in Narrative-Centered Learning Environments

Jonathan P. Rowe, Lucy R. Shores, Bradford W. Mott, James C. Lester, *North Carolina State University, Raleigh, NC, USA*
{jprowe, lrshores, bwmott, lester}@ncsu.edu

Abstract. A key promise of narrative-centered learning environments is the ability to make learning engaging. However, there is concern that learning and engagement may be at odds in these game-based learning environments. This view suggests that, on the one hand, students interacting with a game-based learning environment may be engaged but unlikely to learn, while on the other hand, traditional learning technologies may promote deep learning but provide limited engagement. This paper presents findings from a study with human participants that challenges the view that engagement and learning need be opposed. A study was conducted with 153 middle school students interacting with a narrative-centered learning environment. Rather than finding an oppositional relationship between learning and engagement, the study found a strong positive relationship between learning outcomes, in-game problem solving and increased engagement. Furthermore, the relationship between learning outcomes and engagement held even when controlling for students' background knowledge and game-playing experience. Additional analyses revealed that males tended to report significantly greater presence in the virtual environment than girls, and students with more game-playing experience reported significantly greater presence in the virtual environment than students with minimal game-playing experience. Follow up analyses suggested that differences in presence may be more strongly associated with game-playing experience than gender.

Keywords. Narrative-centered learning environments, game-based learning, engagement, situational interest, presence

INTRODUCTION

Narrative-centered learning environments show significant potential for providing engaging learning experiences that are tailored to individual students. By leveraging the motivational characteristics of narrative and games, along with the adaptive pedagogy of intelligent tutoring systems, narrative-centered learning environments offer a promising platform for students to acquire enhanced problem solving, strategic and analytical thinking, decision making, and other twenty-first century skills (Gee, 2003; Shaffer, 2006). As an active and growing area of research, narrative-centered learning environments are under investigation in a range of domains, including language learning (Johnson, 2007), anti-bullying education (Aylett, Louchart, Dias, Paiva, & Vala, 2005), intercultural negotiation training (Kim, et al., 2009) and middle school science (Ketelhut, 2007).

Despite the ITS community's growing interest in narrative-centered learning environments, there is concern that the narrative and gameplay elements of these systems may not contribute to improved learning outcomes. This concern extends to educational games and game-based learning environments

more broadly. The belief stems in part from a view that gains in engagement achieved by game technologies are primarily diversionary (Hallinen, Walker, Wylie, Ogan, & Jones, 2009; Rai, Heffernan, Gobert, & Beck, 2009). In a recent article, Mayer and Johnson warn that the "entertainment features of games may distract the player from the academic content of the game, and reduce the players' efforts to process the material deeply" (2010, p. 8). These views suggest that while students may become engaged in the rich virtual environments or compelling characters provided by game technologies, the reasons for engagement are tangential to learning (Rowe, McQuiggan, Robison, & Lester, 2009). In this view, there is a tradeoff between learning and engagement, suggesting that on the one hand, students interacting with a game-based learning environment may be engaged but unlikely to learn, and on the other hand, traditional learning technologies may promote deep learning but provide limited engagement (McNamara, Jackson, & Graesser, 2009; Prensky, 2001; Rai, Heffernan, Gobert, & Beck, 2009). In the worst instances, educational games may be neither entertaining nor substantive, combining negative characteristics of both instructional interventions and entertainment-oriented games (Papert, 1998).

This paper challenges the above view by presenting findings from an empirical study investigating the relationships between learning, problem solving and engagement in a narrative-centered learning environment. This work considers a number of factors hypothesized to be associated with engagement, including presence, situational interest, avoidance of "gaming the system," and problem-solving efficiency. Findings are presented from a study with 153 eighth-grade students interacting with CRYSTAL ISLAND, a narrative-centered learning environment for middle school microbiology. Results show that students who experienced higher levels of engagement during interactions with the CRYSTAL ISLAND environment achieved improved learning outcomes and improved in-game problem-solving performance. Notably, the learning results are independent of students' prior microbiology knowledge and game playing experience.

This paper is organized as follows. The next section provides background on narrative-centered learning environments and engagement. Afterward, details of the CRYSTAL ISLAND learning environment are presented. The following two sections describe a study that was conducted with middle school students interacting with the CRYSTAL ISLAND environment, and empirical results from the study. A discussion of the findings and design implications follows. The paper concludes with directions for future work.

BACKGROUND

Narrative-centered learning environments are a class of game-based learning environments that contextualize educational content and problem solving with interactive story scenarios. Multi-user virtual environments such as Quest Atlantis (Barab et al., 2007) and River City (Ketelhut, 2007) use rich narrative settings to contextualize inquiry-based science learning scenarios with strong social and collaborative elements. Other work has utilized interactive narrative generation and agent behavior planning to foster adaptive narrative experiences (Aylett, Louchart, Dias, Paiva, & Vala, 2005; Si, Marsella, & Pynadath, 2005; Traum, Marsella, Gratch, Lee, & Hartholt, 2008). Recently, narrative-centered learning environments have begun to directly incorporate intelligent tutoring facilities, which provide coaching, feedback, and reflection support that is tailored to individual students (Kim, et al., 2009; Thomas & Young, 2009). A key motivation for this line of work is the development of systems that simultaneously promote deep learning and high engagement.

Despite recognition of the importance of engagement in learning and software applications, there has historically been little consensus about its proper definition or constituent parts. In a comprehensive review of the theoretical underpinnings of engagement, O'Brien and Toms propose the following definition: "a quality of user experiences…that is characterized by attributes of challenge, aesthetic and sensory appeal, feedback, novelty, interactivity, perceived control and time, awareness, motivation, interest, and affect" (2008, p. 949). They proceed by presenting a four-stage process model: *point of engagement*, *sustained engagement*, *disengagement*, and *reengagement*. Each of the four stages is associated with a subset of the identified attributes, offering both a conceptual and operational view of the engagement process. However, important questions remain about the relationships between engagement's attributes, its various stages, and different contexts. For example, aesthetics may not be an important component of engagement when solving a complex math problem, but it may be very important during an interaction with a video game. Investigating relevant attributes that can be measured before, during, and after interactions with narrative-centered learning environments is a major component of the current work.

For years, devising techniques for detecting and measuring student engagement has been an important area of investigation within the ITS community (Beck, 2005; Walonoski & Heffernan, 2006; de Vicente & Pain, 2002). A number of techniques have been proposed to assess attributes that are related to engagement, such as student motivation (Johns & Woolf, 2006; de Vicente & Pain, 2002) and affective states such as flow (D'Mello, Picard, & Graesser, 2007). Other work has sought to devise automated models for detecting symptoms of disengagement, namely, off-task behavior (Baker, Corbett, Koedinger, & Wagner, 2004; Walonoski & Heffernan, 2006). One type of student behavior that has been the subject of increasing attention is *gaming the system*, where students exploit elements of a learning environment interface to progress through a lesson without having mastered the associated content (Baker, Corbett, Koedinger, & Wagner, 2004). Baker, Corbett, and Koedinger identify two types of gaming the system: harmful gaming and non-harmful gaming (2004). This distinction comes from the idea that some students game the system because they find a problem difficult, whereas other students game on easy problems in order to progress through a lesson more quickly. Disengagement in narrative-centered learning environments may follow a similar pattern. Some students may disengage because they believe a narrative scenario is challenging or uninteresting; other students may go off-task because the scenario is too simple or because prior game-playing experience motivates them to explore freely rather than advance the narrative.

Engagement in narrative-centered learning environments can take several forms, including engagement in the learning scenario and engagement in tangential or aesthetic elements of the virtual environment (Rowe, McQuiggan, Robison, & Lester, 2009). Narrative-centered learning environments often provide vast interactive environments, realistic physics, and engaging characters, which may risk introducing *seductive details* into learning experiences (Harp & Mayer, 1998). Seductive details have the potential to distract, disrupt, or divert students' attention from pedagogical objectives and reduce students' time-on-task.

Investigating engagement in narrative-centered learning environments requires a multi-faceted approach. This includes using subjective and objective measures, as well as online and post hoc measures, that assess multiple dimensions of narrative-centered learning experiences. For example, presence measures are designed to assess the fidelity and authenticity of an interactive virtual environment (Witmer & Singer, 1998), which is one component of narrative-centered learning experiences. Measures of situational interest can assess how appealing a narrative-centered learning experience is to students (Schraw, 1997), and they encapsulate both game-related and narrative-related

contributors. A student's efficiency in advancing a plot, as well as avoidance of gaming the system behaviors, are partial online indicators of how engaged the student is in a narrative-centered learning scenario. There are numerous additional components of engagement, as indicated by O'Brien and Toms (2008). The work presented in this paper focuses on situational interest, presence, and a researcher-generated measure that quantifies students' narrative efficiency and deliberative problem solving behaviors. These measures are considered in order to assess multiple components of engagement in narrative-centered learning. While additional measures would be necessary to achieve a comprehensive view of engagement, such an investigation remains for future work.

Situational interest is characterized by varying lengths of concentrated attention coupled with affective reaction activated during a particular time period by certain environmental stimuli (Ainley, Hidi, & Berndorff, 2002; Hidi, 1990). Studies have shown that situational interest directed towards an instructional task can influence cognitive performance (Schiefele, 1996) and facilitate deeper learning (Wade, Buxton, & Kelly, 1999). Also, learning tasks and environments that yield significant situational interest have been shown to benefit students who have previously been disengaged in similar learning activities (Hidi & Harackiewicz, 2000). However, situational interest is not exclusive to learning tasks; game design and adaptive scaffolding should encourage interest in on-task actions, rather than interest in purely aesthetic or gameplay features of narrative-centered learning environments (Rowe, McQuiggan, Robison, & Lester, 2009).

Presence contributes to the goal of transparency in technology-mediated interactions (Norman, 1998). Although there has been substantial debate on formal definitions, there is a general consensus that *presence* describes a user's sense of "being there" when interacting with a mediated environment (Schubert, Friedmann, & Regenbrecht, 1999). Presence has been alternatively defined as "the subjective experience of being in one place or environment, even when one is physically situated in another" (Witmer & Singer, 1998). It is related to the sense of transportation into a story, which is an important contributor to the engaging quality of narratives. Presence is distinguished from related concepts such as immersion and involvement. *Immersion* generally refers to the extent and nature of technology-provided sensory stimuli; it is often associated with the pervasiveness and fidelity of visual, auditory, olfactory, and tactile inputs (Schubert, Friedmann, & Regenbrecht, 1999). *Involvement* refers to the degree of attention and meaning devoted to some set of stimuli (Witmer & Singer, 1998).

CRYSTAL ISLAND

Now in its third major iteration, CRYSTAL ISLAND is a narrative-centered learning environment built on Valve Software's Source™ engine, the 3D game platform for Half-Life 2. The curriculum underlying CRYSTAL ISLAND's mystery narrative is derived from the North Carolina state standard course of study for eighth-grade microbiology. The environment is designed as a supplement to classroom instruction, and it blends elements of both inquiry learning and direct instruction. Due to its capacity to teach and engage students, CRYSTAL ISLAND also offers promise as a tool for preparation for future learning (Schwartz & Bransford, 1998). Over the past several years, CRYSTAL ISLAND has served as a platform for investigating a range of artificial intelligence technologies for dynamically supporting students' learning experiences. This includes work on combined narrative and tutorial planning (Mott & Lester, 2006a), narrative director agents (Mott & Lester, 2006b), student modeling (Mott, Lee, & Lester, 2006; Lee, Mott, & Lester, 2010; Rowe & Lester, 2010), archetype-driven
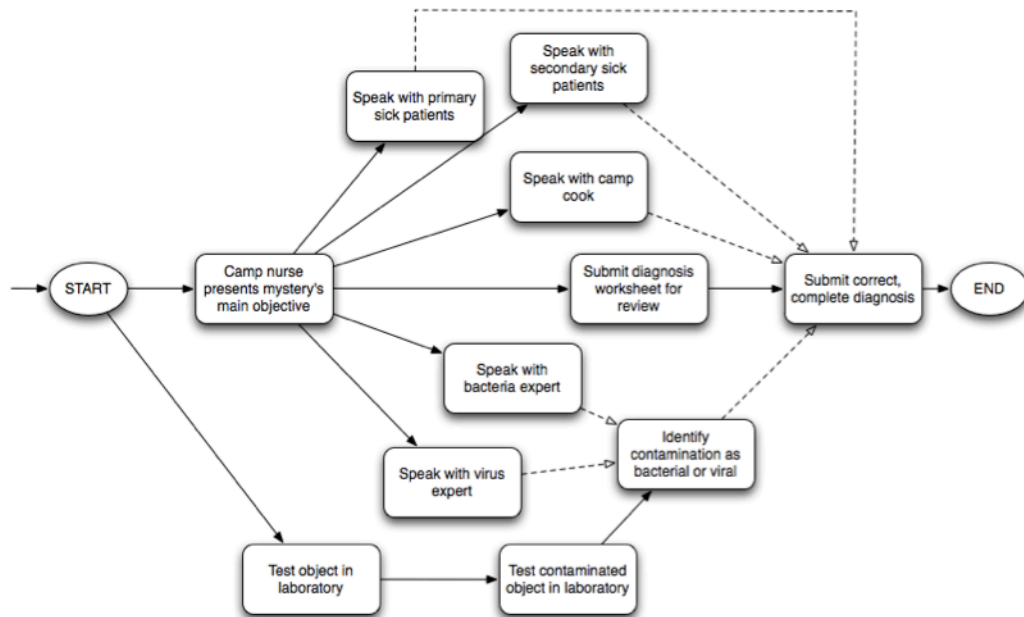
Fig. 1. Partial order among goals in CRYSTAL ISLAND's narrative scenario. Square boxes represent the eleven goals to be completed. Solid arrows indicate precondition relationships between goals. Dotted arrows indicate soft precondition relationships, where completion of the source node's goal contributes to completion of the destination node's goal, but their relative ordering is not strictly enforced.

models of character dialogue (Rowe, Ha, & Lester, 2008), empathetic character behavior models (Robison, McQuiggan, & Lester, 2009), and affect recognition models (McQuiggan, Mott, & Lester, 2008). The environment has also been the subject of extensive empirical investigations of student learning (McQuiggan, Rowe, Lee, & Lester, 2008) and presence (McQuiggan, Rowe, & Lester, 2008), with results informing the design and revision of successive iterations of the system.

CRYSTAL ISLAND's premise involves a mysterious illness that is afflicting a research team stationed on a remote island. The student plays the role of a visitor who recently arrived on the island in order to see her sick father; however, the student gets drawn into a mission to save the entire research team from the spreading outbreak. The student explores the research camp from a first-person viewpoint and manipulates virtual objects, converses with characters, and uses lab equipment and other resources to solve the mystery. The mystery is solved by completing a series of partially ordered goals that uncover details about the spreading infection. The narrative's eleven goals are presented in Figure 1, and are elaborated below.

The scenario begins with the student's arrival at the research camp. A group of small buildings is visible a short distance from the camp entrance. The student approaches the first building, an infirmary, where several sick patients and a camp nurse are lingering. A conversation with the nurse is initiated when the student approaches the character and clicks the mouse. The nurse explains that an unidentified illness is spreading through the camp and asks for the student's help to diagnose the disease (Figure 2). She advises the student to use an in-game diagnosis worksheet in order to record
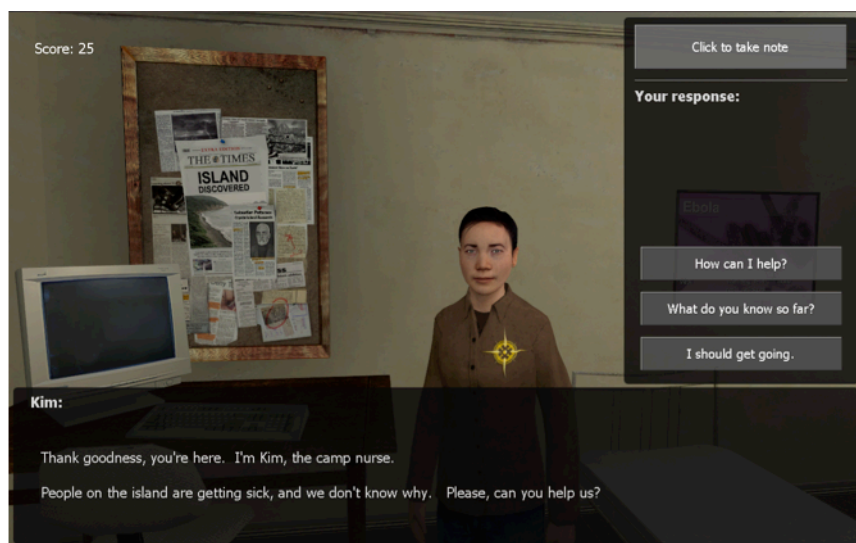
Fig. 2. Conversation with CRYSTAL ISLAND's camp nurse.

her findings, hypotheses, and final diagnosis (Figure 3). This worksheet is designed to scaffold the student's problem solving process, as well as provide a space for the student to offload any findings gathered about the illness. The conversation with the nurse takes place through a combination of multimodal character dialogue—spoken language, gesture, facial expression, and text—and student dialogue menu selections. All character dialogue is provided by voice actors and follows a deterministic branching structure.

After speaking with the nurse, the student has several options for investigating the illness. Inside the infirmary, the student can talk to several sick patients lying on medical cots (Figure 4). Clues about the team members' symptoms and recent eating habits can be discussed and recorded in the diagnosis worksheet or as free-form notes using an in-game smartphone. Alternatively, the student can move to the camp's dining hall to speak with the camp cook. The cook describes the types of food that the team has recently been eating and provides clues about which items warrant closer investigation. In addition to learning about the sick team members, the student has multiple options for gathering information about different disease-causing agents. For example, the student can walk to the camp's living quarters where she will encounter a pair of virtual scientists willing to answer questions about viruses and bacteria, respectively. The student can also learn more about pathogens by viewing posters hanging inside of the camp's buildings (Figure 4) or reading books located in a small library. In this way, the student can gather information about relevant microbiology concepts using resources that are presented in multiple formats. After a conversation with a virtual character, the student may receive a phone call from the camp nurse (using the in-game smartphone) and be asked to answer a short series of multiple-choice questions about microbiology concepts. Kim prefaces the questions by explaining that she is trying to determine the student's progress in solving the mystery; the questions are designed to assess what material the student has retained from the character interactions.

Beyond gathering information from virtual scientists and other instructional resources, the student can conduct tests on food objects using the laboratory's testing equipment (Figure 5). The student encounters food items in the dining hall and laboratory, and she can test the items for pathogenic

Fig. 3. Diagnosis worksheet in which students record their findings in CRYSTAL ISLAND.

contaminants at any point during the learning interaction. For each test, the student must specify the type of test she wishes to conduct and select a justification for that test. A limited number of tests are allocated to the student at the start of the scenario, but additional tests can be earned by answering microbiology multiple-choice questions. Therefore, if a student squanders her available tests by using a haphazard problem-solving strategy, she must demonstrate her understanding of microbiology concepts in order to continue advancing the story.

After running several tests, the student discovers that the sick team members have been consuming contaminated milk. Upon arriving at this finding, the student is instructed to see the lab technician, Elise, for a closer look. The screen momentarily fades to black to indicate elapsing time, and Elise returns with an image of the contaminated specimen, which she explains was taken using a microscope. At this point, the student is presented with a labelling exercise where she must specify the identity and parts of a bacteria cell. After successfully completing this activity, the student can use the camp's books and posters in order to investigate bacterial diseases that are associated with symptoms matching those reported by the sick team members. The student enters her findings and hypotheses into the diagnosis worksheet. Once she has narrowed down a diagnosis and recommended treatment, as well as entered them in her worksheet, the student returns to the infirmary in order to inform the camp nurse. If the student's diagnosis (recorded in the worksheet) is incorrect, the nurse identifies the error and recommends that the student keep working. If the student has correctly diagnosed the illness and specified an appropriate treatment, the mystery is solved.

Fig. 4. Sick patients in the CRYSTAL ISLAND infirmary surrounded
by microbiology-themed posters.

In order to characterize engagement during interactions with CRYSTAL ISLAND, it is useful to identify ways that disengagement surfaces in the environment. For example, students who are inactive in the virtual environment because of distraction or boredom are clearly disengaged. Off-task behaviors that occur within the software are also symptoms of disengagement. A student could be considered off-task if she is interacting extensively with an in-game object that is not relevant to the illness scenario (e.g., a shipping crate or a recycling bin). A student could also be considered off-task if they are manipulating a task-related object in an unrelated location (e.g., dropping a tea kettle on top of a sick character in the infirmary). Students who climb on trees or buildings would be considered off-task. Additionally, a student who loiters in a non-learning oriented location for an excessive duration (e.g., four minutes at the waterfall near the camp edge) would also likely be disengaged.
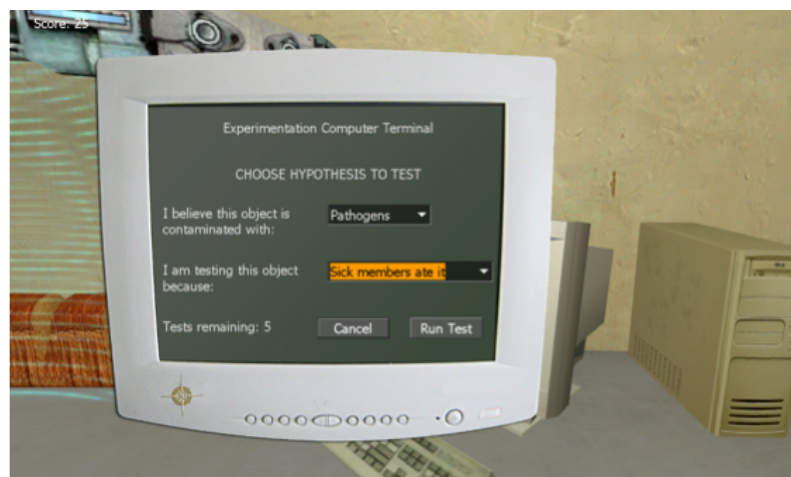

Fig. 5. Testing device in CRYSTAL ISLAND's camp laboratory.

EMPIRICAL STUDY

An experiment involving human participants using CRYSTAL ISLAND was conducted with the entire eighth grade population of a North Carolina middle school. The primary goal of the experiment was to investigate the impact of different coaching techniques on learning and engagement in the CRYSTAL ISLAND narrative-centered learning environment. The coaching included several prompts that were designed to direct students toward in-game learning resources as well as self-regulatory prompts. However, no condition effects were observed for either learning or engagement. This paper's findings stem from a secondary analysis of the data, which considers the experiment's conditions as a whole.

Participants

A total of 153 eighth grade students ranging in age from 12 to 15 ($M = 13.3$, $SD = 0.48$) interacted with the CRYSTAL ISLAND environment during the study. Sixteen of the participants were eliminated due to incomplete data or prior experience with an earlier version of CRYSTAL ISLAND. Among the remaining students, 77 were male and 60 were female. Approximately 3% of the participants were American Indian or Alaska Native, 2% were Asian, 32% were African American, 13% were Hispanic or Latino, and 50% were White. The participants had not yet been exposed to the microbiology curriculum unit of the North Carolina state standard course of study in their regular classes, and therefore had minimal prior experience with CRYSTAL ISLAND's microbiology content.

Materials and Apparatus

Students completed an online demographic survey and CRYSTAL ISLAND curriculum test prior to the intervention. The curriculum test consisted of 16 multiple-choice questions created by an interdisciplinary team of researchers. The test consisted of eight factual and eight application questions assessing students' knowledge of pathogens, select diseases, and the scientific method.

Post-experiment materials were completed immediately following the CRYSTAL ISLAND intervention. Included in these materials were the same curriculum test used in the pre-experiment, a variation of the Perceived Interest Questionnaire (Schraw, 1997), and the Presence Questionnaire (Witmer & Singer, 1998). The interest scale was adapted from measures used by Schraw to examine within-subject relationships with learning outcomes (Schraw, 1997). The measure consists of ten Likert items measuring students' situational interest related to CRYSTAL ISLAND (see Appendix). The Presence Questionnaire (PQ) is a validated measure containing several subscales, including involvement/control, naturalism of experience and quality of interface (Witmer & Singer, 1998). The natural subscale is intended to assess the student's perception of the virtual environment's consistency with reality, in terms of locomotion and nature of the interaction. The interface quality subscale indicates how seamlessly the control and display devices are integrated into the interactive experience. Example items include the following: "How compelling was your sense of moving around inside the virtual environment," "How much did your experiences in the virtual environment seem consistent with your real-world experiences," and "How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?"

In addition to pre- and post-experiment questionnaires, two additional measures were calculated based on students' in-game actions in order to assess student problem-solving and engagement. The measure to assess problem-solving performance was *number of goals completed*, and was calculated

Table 1
Point values for calculation of final game score.

Action	Points (pts)
Overall Mystery Solution	
Correct Solution	500 pts
Solution Efficiency	(7500 / elapsed time) pts
Incorrect Solution Attempt	-100 pts
In-game Quiz Questions	
First Attempt Correct	25 pts
Second Attempt Correct	10 pts
Second Attempt Incorrect	-10 pts
Object Contaminant Testing	
Test Milk for Pathogens	200 pts
Incorrect Object	-10 pts
Incorrect Contaminant	-25 pts
Character Interactions	
Talk to Kim	(25 / elapsed time) pts
Talk to Teresa	(50 / elapsed time) pts
Talk to Ford	(125 / elapsed time) pts
Talk to Robert	(125 / elapsed time) pts
Talk to Quentin	(125 / elapsed time) pts
Pathogen Labeling Activities	
Correct Answer	10 pts
Incorrect Answer	-10 pts
Total Maximum Points	**≈ 1665 pts**

by counting the number of in-game goals students completed during the scenario. The scenario featured eleven possible goals, although not all of them were mandatory for solving the mystery. The open-ended nature of CRYSTAL ISLAND's mystery imposes a partial ordering among the goals, and completion of all eleven is considered to be an important component of effective problem solving in the environment. The eleven goals and their partial ordering are presented in Figure 1. This measure treats all goals equivalently; none of the goals are weighted differently than others.

Another in-game measure used as an indicator of student engagement was *final game score*, which is a numerical score calculated by the CRYSTAL ISLAND software to assess students' progress and efficiency in completing the science mystery. Students could view their scores in the upper left corner of their screens throughout their interactions with the software. The final game score consisted of a weighted sum of gameplay sub-scores, and incorporated time taken to accomplish important narrative goals, students' ability to demonstrate microbiology content knowledge, and evidence of careful hypothesis formulation. Students were penalized for any attempt to "game the system" by repeatedly submitting incorrect diagnoses to the camp nurse or guessing on content knowledge quizzes. Details of final game score's calculation are shown in Table 1.

In the current work, final game score is used to approximate an online measure of engagement because it quantifies the degree to which students exercised learning behaviors that are associated with engaged problem solving in CRYSTAL ISLAND. While the measure does incorporate non-engagement

related factors, such as demonstration of microbiology knowledge, these factors are lightly weighted in the overall calculation. It is not proposed that final game score measures engagement directly, but rather that it is correlated with engagement and provides insight into students' learning experiences beyond what is available from post hoc questionnaires. As an objective online measure characterizing students' gameplay characteristics in the CRYSTAL ISLAND environment, final game score provides a tool for distinguishing students who were disengaged in the scenario (as evidenced by inefficient goal completion or repeated attempts to guess the solution) and students who were engaged in the problem solving task.

Participant Procedure

Participants entered the experiment room having completed the majority of pre-test materials one week prior to the intervention. Students were initially provided general details about the CRYSTAL ISLAND mystery and game controls during an introductory presentation by a researcher. After the presentation, students completed the remaining pre-test materials and received several CRYSTAL ISLAND supplementary documents. These materials consisted of a CRYSTAL ISLAND backstory and task description, a character handout, a map of the island, and an explanation of the game's controls.

Participants were given 60 minutes to work on solving the mystery. Immediately after solving CRYSTAL ISLAND's science mystery, or after 60 minutes of interaction, participants completed the post-experiment questionnaires. Completion of post-experiment materials took no longer than 30 minutes for participants. In total, sessions lasted up-to 120 minutes.

RESULTS

An investigation of learning found that on average, students answered 2.35 (SD = 2.75) more questions correctly on the post-test than they did on the pre-test. Matched pairs t-tests (comparing post-test to pre-test scores) indicated that students' learning gains were significant, $t(149)$ = 10.5, $p < .001$.

Learning and Engagement

Examining factors believed to reflect engagement and students' understanding of the curriculum, Pearson correlations indicated significant relationships between microbiology background knowledge and presence (r = .17, $p < .05$) and final score (r = .28, $p < .01$). Similar relationships were found between microbiology post-test scores and presence (r = .29, $p < .01$), final score (r = .45, $p < .01$), and situational interest (r = .24, $p < .01$). To more closely investigate the relationships between learning and engagement, additional analyses controlling for background knowledge were conducted.

A partial correlation controlling for pre-test score found significant relationships between microbiology post-test scores and two of our engagement measures, presence (r = .25, $p < .01$), and final game score (r = .38, $p < .01$). The same type of analysis also found a borderline significant relationship between situational interest and post-test score (r = .15, $p < 0.1$). Offering further evidence for a connection between learning and engagement in CRYSTAL ISLAND, a linear regression indicated that microbiology background knowledge, presence, and final game score were all

significant predictors of performance on the microbiology post-test, and the model as a whole was significant ($R^2 = .33$, $F(3, 143) = 23.5$, $p < .001$; see Table 2).

As a supplement to these findings, further analyses were conducted to determine whether similar relationships held for the involved/control subscale of the Presence Questionnaire (PQ), which provides a more specific measure of involvement in the environment. A partial correlation controlling for microbiology background knowledge revealed significant relationships between the involved/control subscale and final score ($r = .38$, $p < .01$), situational interest ($r = .18$, $p < .05$), and microbiology post-test performance ($r = .33$, $p < .01$).

Table 2
Regression results predicting students' microbiology post-test performance.

Predictor	B	SE	β
Pre-Test	.46**	.09	.33**
Presence	.03*	.01	.15*
Final Game Score	.01**	.00	.31**
R^2		.33	

Note: ** - $p < .01$; * - $p < .05$

Problem Solving and Engagement

Analyses were conducted to investigate the relationship between the engagement-related measures and in-game problem-solving performance, as determined by the number of goals completed measure ($M = 9.26$, $SD = 1.91$). In this subsection, only presence and situational interest are investigated as engagement-related variables. These analyses do not consider final game score, because its calculation is based in part on how efficiently students completed a subset of CRYSTAL ISLAND's eleven goals.

Pearson correlations revealed significant relationships between number of goals completed and microbiology background knowledge ($r = .17$, $p < .05$), microbiology post-test scores ($r = .47$, $p < .01$), situational interest ($r = .21$, $p < .05$), and presence ($r = .27$, $p < .01$). Similarly, a partial correlation analysis controlling for microbiology background knowledge found significant relationships between number of goals completed and microbiology post-test performance ($r = .40$, $p < .01$) and presence ($r = .24$, $p < .01$).

To further investigate the relationship between in-game problem solving and engagement, a hierarchical regression analysis was conducted controlling for microbiology background knowledge (see Table 3). In order to predict microbiology post-test performance, microbiology pre-test score was entered into the first block while situational interest, presence, and number of goals completed were entered into the second block. Microbiology background knowledge was found to be a significant predictor of microbiology post-test performance ($F(1, 130) = 28.6$, $p < .01$) and was responsible for 18% of the variance. However, the second model was also found to significantly predict microbiology post-test performance ($F(4, 127) = 16.9$, $p < .01$). Microbiology pre-test score, number of goals completed, and presence were all significant predictors of microbiology post-test performance, and

accounted for 35% of post-test score's variance. Situational interest was not observed to be a significant predictor.

Table 3
Hierarchical regression results predicting students' microbiology post-test performance.

Predictor	Model 1			Model 2		
	B	SE	β	B	SE	β
Pre-Test	.61**	.10	.44**	.46**	.11	.32**
Presence				.03*	.02	.16*
Goals Completed				.54**	.13	.33**
Situational Interest				.02	.03	.05
R^2		.18			.35	

Note: ** - $p < .01$; * - $p < .05$

Gender and Engagement

Analyses were conducted to determine whether particular subpopulations experienced different levels of engagement while interacting with the CRYSTAL ISLAND environment. An independent samples t-test analyzing the relationship between gender and presence found that males reported greater presence in the environment than females ($t(139) = 3.01$, $p < .01$). Similar results were found for the involved/control subscale of the PQ. An independent samples t-test found that males tended to feel significantly more involved/control when interacting with CRYSTAL ISLAND than females ($t(140) = 2.96$, $p < .01$). Males also tended to rate the interface quality more highly ($t(140) = 1.97$, $p < .01$). However, no gender effect was found on the PQ's natural subscale. Table 4 displays raw scores, by gender, for each of the content knowledge, situational interest, final game score, and presence measures.

Significant differences were observed between genders for game-playing demographics. Males reported significantly higher ratings for self-perceived game-playing skill ($F(1, 143) = 57.5$, $p < .001$) and reported playing games more frequently ($F(1, 143) = 60.2$, $p < .001$) than females. Although males tended to feel more present in CRYSTAL ISLAND, an analysis of covariance (ANCOVA) controlling for game-playing frequency found no significant effect of gender on presence ($F(1, 138) = 2.01$, $p = .158$). Significant differences were not found between genders for situational interest or final game score.

A linear regression considering only the female population yielded a significant model for predicting microbiology post-test performance ($R^2 = .25$, $F(2, 62) = 10.1$, $p < .01$), but only microbiology background knowledge and final score were significant predictors, not presence.

Table 4

Raw Scores by gender, gameplay frequency, and gameplay skill on pre-test, post-test, situational interest, final game score, and presence. LFG – Low game-playing frequency. MGF – Medium game-playing frequency. HGF – High game-playing frequency. LGS – Low game-playing skill. MGS – Medium game-playing skill. HGS – High game-playing skill.

Group	Microbiology Pre-Test	Microbiology Post-Test	Situational Interest	Final Game Score	Presence
Males (N = 78)	6.37 (2.23)	8.60 (3.03)	31.8 (8.73)	328.9 (555.6)	89.5 (16.4)
Females (N = 64)	6.31 (1.77)	8.62 (2.94)	31.4 (8.37)	304.1 (462.6)	82.3 (15.6)
LGF (N = 39)	6.26 (1.57)	8.51 (3.41)	30.2 (9.32)	305.6 (542.9)	80.6 (17.9)
MGF (N = 50)	6.50 (2.10)	8.80 (2.48)	31.1 (8.07)	301.2 (467.0)	85.9 (13.3)
HGF (N = 53)	6.28 (2.29)	8.74 (2.88)	33.3 (8.44)	342.4 (543.0)	91.4 (15.4)
LGS (N = 28)	5.96 (1.53)	8.39 (3.25)	29.4 (9.02)	288.8 (493.4)	78.9 (15.7)
MGS (N = 42)	6.43 (1.99)	9.05 (2.94)	32.1 (7.59)	332.8 (526.2)	87.5 (15.2)
HGS (N = 72)	6.46 (2.23)	8.61 (2.74)	32.4 (8.64)	320.3 (521.6)	89.0 (15.8)
Total	6.34 (2.02)	8.61 (2.98)	31.6 (8.54)	317.8 (514.2)	82.6 (16.4)

Gameplay Experience and Engagement

A regression analysis was conducted to examine the simultaneous contributions of game-playing frequency, microbiology background knowledge, presence, and final game score on predicting microbiology post-test scores. The overall model was significant ($R^2 = .33$, $F(4, 136) = 16.5$, $p < .01$), but only microbiology background knowledge, presence, and final game score were significant predictors of post-test performance, not game-playing frequency. A similar regression analysis was conducted to examine the contributions of self-assessed game-playing skill, microbiology background knowledge, presence, and final game score on microbiology post-test scores. The overall model was significant ($R^2 = .33$, $F(4, 136) = 16.8$, $p < .01$), but again only microbiology background knowledge, presence, and final score were significant predictors, not self-assessed game-playing skill.

In order to investigate the relationships between game-playing experience and engagement, data about students' game-playing frequency and game-playing skill was divided into three groups. For the game-playing frequency data, students who indicated that they played games *rarely* or *not at all* were assigned to the low group (LGF), students who indicated that they played games *occasionally* were placed in the medium group (MGF), and students who indicated that they played games *frequently* or *very frequently* were placed in the high group (HGF). For the game-playing skill data, students who indicated that they were *not at all* skilled or had *limited skills* were assigned to the low group (LGS), students who indicated that they had *average* skills were placed in the medium group (MGS), and students who indicated that they were *skilled* or *very skilled* at playing games were placed in the high group (HGS). The number of students in each of the game-playing frequency and game-playing skill groups is shown in Table 4. The table also displays raw scores, by game-playing frequency and skill, for each of the content knowledge, situational interest, final game score and presence measures.

An analysis of variance (ANOVA) found significant differences between game-playing frequency group and presence ($F(2, 138) = 5.56, p < .01$). Post-hoc analysis found students in the low group reported feeling significantly less present than students in the high group ($p < .01$). No effect of game-playing frequency was observed on number of goals completed. A similar ANOVA found significant differences between perceived game-playing skill and presence ($F(2, 138) = 4.31, p < .05$). Post-hoc analysis found students in the low group felt significantly less present than students in the high game-playing skill group ($p < .05$). An ANOVA also found significant differences in number of goals completed among the game-playing skill groups ($F(2, 124) = 3.46, p < .05$). Post-hoc analysis found students in the low group completed significantly fewer goals than students in the high group ($p < .05$). While controlling for the number of goals completed, an analysis of covariance (ANCOVA) found no differences between perceived game-playing skill and reported presence.

DISCUSSION

The findings indicate that student engagement with the CRYSTAL ISLAND environment was associated with improved learning outcomes and in-game problem solving. Results showed a significant relationship between students' pre-test scores and presence, as well as between pre-test scores and final game scores. This suggests that students who demonstrated greater prior content knowledge tended to become more engaged with the narrative environment. However, all three measures for engagement—presence, situational interest, and final game score—were found to be significantly associated with post-test score, independent of pre-test score. Additionally, presence and situational interest were found to be significantly associated with in-game problem-solving performance, independent of pre-test score.

These findings suggest that students who were more engaged with the CRYSTAL ISLAND narrative environment tended to experience greater learning gains and increased problem solving performance, regardless of prior knowledge. The findings contrast with perspectives that place engagement and learning at odds with one another in narrative-centered learning environments. Further, analyses found no relationships between game-playing experience and learning. This finding suggests that both gamers and non-gamers who were engaged in the narrative-centered learning experience achieved improved learning outcomes. Students can be productively engaged in a narrative-centered learning environment, and this relationship is independent of prior knowledge or game-playing experience.

There are several possible explanations for the presence findings. One explanation is that the association between presence and learning is motivational in nature; students who experience greater presence in the environment are also more motivated to complete the scenario, devote greater focus and attention to the learning interaction, and engage in deeper cognitive processing. An alternative explanation stems from the findings related to game-playing experience. It is plausible that students with less game-playing experience would be less equipped to navigate through the virtual environment and use the game's controls, and would consequently experience decreased presence in the environment. Difficulties with the game's controls could introduce increased cognitive load and consume cognitive resources that might otherwise be dedicated to learning and problem solving. This would potentially explain the associations between game-playing experience and presence, as well as game-playing experience and in-game problem solving. However, more detailed investigation of students' use of the game controls and their rates of gameplay activity in the environment would be necessary to further investigate this possible explanation.

Interesting findings were also observed concerning the effects of gender and game-playing experience on presence. Males tended to be more present during CRYSTAL ISLAND interactions than females. An initial interpretation might be that the game was better designed for males than females. However, a significant effect of game-playing experience was also observed on presence. Furthermore, males tended to have significantly greater game-playing experience than females. An ANCOVA suggested that game-playing experience, not gender, may be the more predominant factor associated with presence.

Measuring Engagement

Given the encouraging findings regarding engagement and learning in narrative-centered learning environments, it is appropriate to reflect on the engagement measures that were chosen. The Presence Questionnaire proved to be useful for assessing CRYSTAL ISLAND's ability to provide students with a sense of transportation into a virtual environment. The measure was a consistent predictor of learning and problem solving performance, and it revealed interesting differences between students of different genders and game-playing experience. In prior work, the Presence Questionnaire has also proven to be useful for assessing the impacts of empathetic characters in narrative-centered learning environments (McQuiggan, Rowe, & Lester, 2008). As a tool for assessing the immersive qualities of virtual environments, the Presence Questionnaire has considerable promise for investigations of narrative-centered learning experiences.

The Perceived Interest Questionnaire was comparatively less consistent in its ability to characterize students' interest in the virtual environment. One possible explanation could be that several of the measure's items were not well matched for the study design. For example, an item such as, "I would play CRYSTAL ISLAND again if I had the chance," may not be applicable to a narrative experience that is designed to be played only once. Furthermore, an item such as "CRYSTAL ISLAND was one of the more interesting games I have played in a long time" may unfairly place the learning environment in competition with big-budget entertainment-focused games that students play at home. The item "CRYSTAL ISLAND was personally relevant to me" was also unlikely to distinguish engaged and non-engaged students. None of the participants had prior experience with the microbiology curriculum used in CRYSTAL ISLAND; therefore, they would not be expected to find the scenario particularly relevant to their own lives. Alternative measures, such as motivation-focused or narrative-

focused scales, may be more appropriate for future work assessing subjective engagement factors that are distinct from presence.

While final game score proved to be a useful predictor of student post-test knowledge, the measure has weaknesses because it conflates factors related to both engagement and content understanding. Rather than suggesting that final game score should be a general measure for engagement, the findings imply that individually investigating the factors that compose final game score will likely yield valuable insights into student engagement and learning in narrative-centered learning environments. Factors that merit further investigation include off-task behavior, rate of narrative activity, and gaming the system.

Designing Narrative-Centered Learning Environments

We hypothesize that carefully designed story and gameplay elements are primary contributors to the synergistic relationship between learning and engagement in narrative-centered learning environments. However, poorly designed story and gameplay elements seem likely to detract from engagement and learning by introducing seductive details and promoting off-task behavior. Additional investigation is needed to determine which elements of narrative-centered learning environments are most closely associated with learning and engagement. These efforts will contribute to the development of models to automatically detect student engagement and learning during narrative-centered learning interactions.

In the case of CRYSTAL ISLAND, tight integration between educational content, narrative, and gameplay is likely a leading factor in the positive relationship between learning and engagement. In order to complete the narrative, students were required to apply knowledge about the microbiology subject matter. Furthermore, students who progressed farther in the narrative were likely to spend more time on microbiology content than students who did not advance through the narrative. Gameplay mechanics, such as students earning additional tests for the laboratory equipment by answering microbiology quiz questions, also nurtured the relationship between content learning and engaged problem solving.

A second likely factor contributing to the relationship between learning and engagement is the design of CRYSTAL ISLAND's narrative. The story is sufficiently detailed to motivate the educational problem-solving task and preserve an authentic and consistent story world, but it is simplistic enough to avoid distracting students from the software's primary learning objectives. Earlier versions of CRYSTAL ISLAND featured more elaborate narratives, but they were abandoned. While detailed narratives may have superficial motivational benefits, extraneous narrative elements can add cognitive load that detracts from students' capacity to learn. For this reason, CRYSTAL ISLAND's narrative is primarily designed to support its learning objectives and prove compelling enough to motivate the student to seek out the narrative's resolution.

With regard to promoting engagement, it is likely that neither the virtual environment nor the narrative played a more important role than the other. Presenting a rich, interactive virtual environment without a story would only hold students' interest for a short amount of time, whereas presenting an interactive mystery scenario without the features of a commercial game environment would likely result in reduced engagement and motivation. CRYSTAL ISLAND was designed to enable gameplay and narrative elements to work in concert to encourage problem solving and learning. The reported findings are consistent with this deliberate design effort. However, it is worth noting that it is possible to devise non-narrative educational games and non-interactive narratives that are

pedagogically effective; the design principles for these types of educational software are likely distinct from those for narrative-centered learning environments.

Future Directions

Extending studies of narrative-centered learning interactions beyond individual sessions is an essential next step for understanding the relationship between engagement and learning in narrative-centered learning environments. Conducting studies spanning multiple sessions, together with classroom integration, is important to assess how engagement can be sustained over time with narrative-centered learning environments, how long-term engagement is related to deep learning and transfer, and whether engagement can impact student attitudes and self-efficacy. Additionally, it will be critical to design appropriate control treatments against which narrative-centered learning environments can be compared. Designing these treatments to ensure that they employ the same content and instructional methods as narrative-centered learning environments presents significant challenges. However, this line of investigation is important for identifying the circumstances under which narrative-centered learning environments are most effective. As discussed above, devising additional subjective and objective measures for engagement beyond those used in this work will also be important for better understanding the impacts of narrative-centered learning environments.

CONCLUSION

Narrative-centered learning environments offer a promising vehicle for delivering experiences that are both effective and engaging. To investigate the hypothesis that learning and engagement need not be in opposition in narrative-centered learning environments, an empirical study was conducted with middle school students interacting with the CRYSTAL ISLAND learning environment. It was found that increased engagement was associated with improved learning outcomes and problem solving, independent of students' prior content knowledge or game-playing experience. Additionally, individual differences in presence were observed to be more likely associated with game-playing experience than gender. As narrative-centered learning environments mature, it will become increasingly important to understand how students can most effectively interact with them, which students stand to benefit most, and what role narrative and game features can play in scaffolding learning and realizing sustained engagement.

ACKNOWLEDGEMENTS

REFERENCES

Ainley, M., Hidi, S., & Berndorff, D. (2002). Interest, learning, and the psychological processes that mediate their relationship. *Journal of Educational Psychology*, *94*(3), 545-561.

Aylett, R. S., Louchart, S., Dias, J., Paiva, A., & Vala, M. (2005). FearNot! An experiment in emergent narrative. In T. Panayiotopoulos, J. Gratch, R. Aylett, D. Ballin, P. Olivier, & T. Rist (Eds.) *Proceedings of the 5th International Conference on Intelligent Virtual Agents* (pp. 305-316). Berlin: Springer-Verlag.

Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004). Off-task behavior in the Cognitive Tutor classroom: When students "game the system." In E. Dykstra-Erickson & M.Tscheligi (Eds.) *Proceedings of the ACM Conference on Human Factors in Computing Systems* (pp. 383-390). New York: ACM.

Baker, R. S., Corbett, A. T., & Koedinger, K. R. (2004). Detecting student misuse of intelligent tutoring systems. In J. Lester, R. M. Vicari, & F. Paraguac (Eds.) *Proceedings of the 7ᵗʰ International Conference on Intelligent Tutoring Systems* (pp. 531-540). Berlin: Springer-Verlag.

Barab, S., Dodge, T., Tuzun, H., Job-Sluder, K., Jackson, C., Arici, A., et al. (2007). The Quest Atlantis project: A socially-responsive play space for learning. In B. E. Shelton & D. Wiley (Eds.), *The Educational Design and Use of Simulation Computer Games* (pp. 159-186). Rotterdam, The Netherlands: Sense Publishers.

Beck, J. E. (2005). Engagement tracing: using response times to model student disengagement. In R. Luckin, K. R. Koedinger & J. Greer (Eds.) *Proceedings of the 13th International Conference on Artificial Intelligence in Education* (pp. 88-95). Amsterdam: IOS Press.

Burton, R. R. & Brown, J. S. (1976). A tutoring and student modelling paradigm for gaming environments. In R. Colman & P. Lorton, Jr. (Eds.) *Proceedings of the ACM SIGCSE-SIGCUE Technical Symposium on Computer Science and Education*, (pp. 236-246). New York: ACM.

D'Mello, S., Picard, R. W., & Graesser, A. (2007). Toward an affect-sensitive AutoTutor. *Special Issue on Intelligent Educational Systems – IEEE Intelligent Systems*, *22*, 53-61. IEEE Computer Society.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.

Hallinen, N., Walker, E., Wylie, R., Ogan, A., & Jones, C. (2009). I was playing when I learned: A narrative game for French aspectual distinctions. In S. Craig & D. Dicheva (Eds.) *Proceedings of the Workshop on Intelligent Educational Games at the 14th International Conference on Artificial Intelligence in Education* (pp. 117-120). Berlin: Springer-Verlag.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*(3), 414-434.

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research*, *60*(4), 549-571.

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research*, *70*(2), 151-179.

Johns, J., & Woolf, B. (2006). A dynamic mixture model to detect student motivation and proficiency. In A. Cohn (Ed.) *Proceedings of the 21st National Conference on Artificial Intelligence* (pp. 163-168). Menlo Park, California: AAAI Press.

Johnson, W. L. (2007). Serious use of a serious game for language learning. In Dimitrova, Mizoguchi, du Boulay, & A. Graesser (Eds.) *Proceedings of the 13th International Conference on Artificial Intelligence in Education* (pp. 67-74). Berlin: Springer-Verlag.

Ketelhut, D. J. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment. *Journal of Science Education and Technology*, *16*, 99-111.

Kim, J., Hill, R. W., Durlach, P., Lane, H. C., Forbell, E., Core, M., Marsella, S., Pynadath, D., Hart, J. (2009). BiLAT: A game-based learning environment for practicing negotiation in a cultural context. *International Journal of Artificial Intelligence in Education*, 19(3), 289-308.

Lee, S., Mott, B., & Lester, J. (2010). Optimizing story-based learning: An investigation of student narrative profiles. In V. Aleven, J. Kay, & J. Mostow (Eds.) *Proceedings of the Tenth International Conference on Intelligent Tutoring Systems* (pp.155-165). Berlin: Springer-Verlag.

Mayer, R. E., & Johnson, C. I. (2010). Adding instructional features that promote learning in a game-like environment. *Journal of Educational Computing Research*, *42*(3), 241-265.

McNamara, D. S., Jackson, G. T., & Graesser, A. (2009). Intelligent tutoring and games (ITaG). In S. Craig & D. Dicheva (Eds.) *Proceedings of the Workshop on Intelligent Educational Games at the 14th International Conference on Artificial Intelligence in Education* (pp. 1-10). Berlin: Springer-Verlag.

McQuiggan, S., Mott, B., & Lester, J. (2008). Modeling self-efficacy in intelligent tutoring systems: An inductive approach. *User Modeling and User-Adapted Interaction, 18*(1-2), 81-123.

McQuiggan, S., Rowe, J., Lee, S. & Lester, J. (2008). Story-based learning: The impact of narrative on learning experiences and outcomes. In B. Woolf, E. Aïmeur, R. Nkambou, & S. Lajoie (Eds.) *Proceedings of the Ninth International Conference on Intelligent Tutoring Systems* (pp.530-539). Berlin: Springer-Verlag.

McQuiggan, S., Rowe, J. & Lester, J. (2008). The effects of empathetic virtual characters on presence in narrative-centered learning environments. In M. Czerwinski, A. Lund, D. Tan (Eds.) *Proceedings of the 2008 SIGCHI Conference on Human Factors in Computing Systems* (pp. 1511-1520). New York: ACM.

Mott, B. & Lester, J. (2006a). U-Director: A decision-theoretic narrative planning architecture for storytelling environments. In E. André, S. Sen, C. Frasson, & J. P. Müller (Eds.) *Proceedings of the Fifth International Conference on Autonomous Agents and Multiagent Systems* (pp. 977-984). New York: ACM.

Mott, B. & Lester, J. (2006b). Narrative-centered tutorial planning for inquiry-based learning environments. In M. Ikeda, K. Ashley, & T. Chan (Eds.) *Proceedings of the Eighth International Conference on Intelligent Tutoring Systems* (pp. 675-684). Berlin: Springer-Verlag.

Mott, B., Lee, S., & Lester, J. (2006). Probabilistic goal recognition in interactive narrative environments. In Y. Gil & R. Mooney (Eds.) *Proceedings of the Twenty-First National Conference on Artificial Intelligence* (pp. 187-192). Menlo Park, California: AAAI Press.

Norman, D. A. (1998). *The invisible computer*. Cambridge, MA: MIT Press.

O'Brien, H. & Toms, E. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, *59*(6), 938-955.

Papert, S. (1998). Does easy do it? Children, games, and learning. *Game Developer, 5*(6), 88.

Prensky, M. (2001). *Digital game-based learning*. New York: McGraw-Hill.

Rai, D., Heffernan, N. T., Gobert, J. D., & Beck, J. E. (2009). Mily's World: Math game involving authentic activities in visual cover story. In S. Craig & D. Dicheva (Eds.) *Proceedings of the Workshop on Intelligent Educational Games at the 14th International Conference on Artificial Intelligence in Education* (pp. 125-128). Berlin: Springer-Verlag.

Robison, J., McQuiggan, S., & Lester, J. (2009). Modeling task-based vs. affect-based feedback behavior in pedagogical agents: An inductive approach. In S. Craig & D. Dicheva (Eds.) *Proceedings of the Fourteenth International Conference on Artificial Intelligence in Education* (pp. 25-32). Berlin: Springer-Verlag.

Rowe, J. & Lester, J. (2010). Modeling user knowledge with dynamic Bayesian networks in interactive narrative environments. In G. M. Youngblood (Ed.) *Proceedings of the Sixth Annual AI and Interactive Digital Entertainment Conference* (pp.57-62). Menlo Park, California: AAAI Press.

Rowe, J., Ha, E. Y., & Lester, J. (2008). Archetype-driven character dialogue generation for interactive narrative. In H. Prendinger, J. Lester, & M. Ishizuka (Eds.) *Proceedings of the Eighth International Conference on Intelligent Virtual Agents* (pp. 45-58). Berlin: Springer-Verlag.

Rowe, J. P., McQuiggan, S. W., Robison, J. L., & Lester, J. C. (2009). Off-task behavior in narrative-centered learning environments. In S. Craig & D. Dicheva (Eds.) *Proceedings of the 14th International Conference on Artificial Intelligence in Education* (pp. 99-106). Berlin: Springer-Verlag.

Schiefele, U. (1996). Topic interest, text representation, and quality of experience. *Contemporary Educational Psychology*, *21*(1), 3-18.

Schraw, G. (1997). Situational Interest in Literary Text. *Contemporary Educational Psychology*, *22*(4), 436-456.

Schubert, T., Friedmann, F., & Regenbrecht, H. (1999). Embodied presence in virtual environments. In R. Paton, *Visual representations and interpretations* (pp. 269-278). London: Springer-Verlag.

Schwartz, D. & Bransford, J. (1998). A Time for Telling. *Cognition and Instruction, 16*(4), 475-522.

Shaffer, D. W. (2006). *How computer games help children learn*. New York: Palgrave Macmillan.

Si, M., Marsella, S. C., & Pynadath, D. V. (2005). Thespian: An architecture for interactive pedagogical drama. In C. K. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Eds.) *Proceedings of the 12th International Conference on Artificial Intelligence in Education* (pp. 595-602). Berlin: Springer-Verlag.

Thomas, J. & Young, R. M. (2009). Using task-based modeling to generate scaffolding in narrative-guided exploratory learning environments. In S. Craig & D. Dicheva (Eds.) *Proceeding of the 14th International Conference on Artificial Intelligence in Education* (pp. 107-114). Berlin: Springer-Verlag.

Traum, D., Marsella, S., Gratch, J., Lee, J., & Hartholt, A. (2008). Multi-party, multi-issue, multi-strategy negotiation for multi-modal virtual agents. In H. Prendinger, J. Lester, & M. Ishizuka (Eds.) *Proceedings of the 8th International Conference on Intelligent Virtual Agents* (pp. 117-130). Berlin: Springer-Verlag.

Wade, S., Buxton, W., & Kelly, M. (1999). Using think-alouds to examine reader-text interest. *Reading Research Quarterly*, *34*(2), 194-216.

Walonoski, J., & Heffernan, N. (2006). Detection and analysis of off-task gaming behavior in intelligent tutoring systems. In M. Ikeda, K. Ashley, & T. Chan (Eds.) *Proceedings of the 8th International Conference on Intelligent Tutoring Systems* (pp. 382-391). Berlin: Springer-Verlag.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225-240.

de Vicente, A., & Pain, H. (2002). Informing the detection of the students' motivational state: an empirical study. In S. Cerri, G. Gouardères, & F. Paraguaçu (Eds.) *Proceedings of the 6th International Conference on Intelligent Tutoring Systems* (pp. 933-943). Berlin: Springer-Verlag.

Appendix

Perceived Interest Scale adapted from (Schraw, 1997) for the CRYSTAL ISLAND study.

1. I thought Crystal Island was very interesting.

2. I'd like to discuss Crystal Island with others at some point.

3. I would play Crystal Island again if I had the chance.

4. I got absorbed playing Crystal Island without trying to.

5. I will probably think about what I learned playing Crystal Island for some time to come.

6. I thought Crystal Island's topic was fascinating.

7. Crystal Island was personally relevant to me.

8. I would like to play more games like Crystal Island in the future.

9. Crystal Island was one of the more interesting games I have played in a long time.

10. Crystal Island really grabbed my attention.