Story-based Learning: The Impact of Narrative on Learning Experiences and Outcomes

Scott W. McQuiggan, Jonathan P. Rowe, Sunyoung Lee, and James C. Lester

Department of Computer Science, North Carolina State University, Raleigh NC 27695
{swmcquig, jprowe, slee7, lester}@ncsu.edu

Abstract. Within the intelligent tutoring systems community, narrative is emerging as an effective medium for contextualizing learning. To date, relatively few empirical studies have been conducted to assess learning in narrative-centered learning environments. In this paper, we investigate the effect of narrative on learning experiences and outcomes. We present results from an experiment conducted with eighth-grade middle school students interacting with a narrative-centered learning environment in the domain of microbiology. The study found that students do exhibit learning gains, that those gains are less than those produced by traditional instructional approaches, but that the motivational benefits of narrative-centered learning with regard to self-efficacy, presence, interest, and perception of control are substantial.

1   Introduction

Narrative is the subject of increasing interest within the intelligent tutoring systems community. Researchers have begun to develop narrative-centered learning environments (NLEs) that combine story contexts and pedagogical support strategies to deliver effective, engaging educational experiences. Contextualizing learning within narrative affords the use of artificial intelligence techniques that tailor narrative and educational content to students' actions, affective states, and abilities. Drawing on an interdisciplinary body of work, including intelligent tutoring systems, embodied conversational agents, and serious games, these environments offer the promise of adaptive, motivating learning experiences to students. NLEs are currently under investigation in a range of domains, including military soft-skills training [7,15], anti-bullying education [1], health intervention education [11], and science learning in microbiology and genetics [12].

By incorporating learning into narrative-based, virtual environments, researchers hope to tap into students' innate facilities for crafting and understanding stories. Contextualized learning experiences are known to encourage regulated learning behavior [14]. Narrative is also well suited to alternative learning paradigms such as guided discovery and inquiry-based learning. Leveraging stories' ability to draw audiences into plots and settings, NLEs can introduce novel perceptual, emotional, and motivational experiences, as well as establish connections between narrative content and pedagogical subject matter in young learners [19]. Further, NLEs can

effectively support the factors shown to contribute to student levels of motivation. Such contextual experiences influence student learning and motivation [8].

There is a strong theoretical foundation and several active projects that support this line of work [1], but there has been limited empirical investigation of learning outcomes within narrative-centered learning environments. Because of the challenges inherent in developing and deploying these types of systems, learning outcome evaluations in previous work has largely been subjective or preliminary in scope. This papers seeks to provide an empirical basis for the evaluation and investigation of NLEs. It presents results from an empirical study conducted with eighth-grade middle school students interacting with an "early generation" NLE, Crystal Island.

2   Related Work

Much of the work on NLEs has focused on developing AI-based approaches that provide rich, adaptive narrative-based learning experiences and respond appropriately to student actions in the environment. FearNot! is a character-driven learning environment for the domain of anti-bullying social education [1]. The environment emphasizes autonomous, highly affective characters that foster empathetic relationships with students, who in turn offer coping suggestions to the victimized virtual character. FearNot! has been the subject of several small- and large-scale studies, although the subjective nature of the domain renders objective, learning-gain results impractical. Carmen's Bright IDEAS seeks to teach health intervention skills to mothers of pediatric cancer patients [11]. The environment combines autonomous characters with director and cinematographic agents in order to provide a dramatic story that satisfies pedagogical goals. Students control the main character's (Carmen's) decisions as she copes with the stresses and problems inherent in caring for an ill child. Carmen's Bright IDEAS has been the subject of clinical trials, but reported results have also been limited.

Intelligent NLEs have recently been developed for military soft-skills training, particularly in leadership and language learning scenarios. IN-TALE is an interactive narrative system that integrates autonomous character behaviors and an Automated Story Director to provide dramatic simulation experiences for social and cultural leadership training [15]. The system draws upon previous work in narrative planning and believable agent behavior to balance narrative coherence and user-agency in the simulation environment. The Tactical Language and Culture Training System (TLCTS) is a story-centric, serious game designed for language learning [7]. TLCTS use a combination of interactive lessons and games to train students in spoken and non-verbal communication, as well relevant cultural knowledge. Over the course of the last several years, TLCTS has transitioned into widespread use by the US military and other groups. However, large-scale, summative empirical results for learning outcomes have not yet been presented for either IN-TALE or TLCTS [6].

Despite the presence of several promising and successful examples of NLEs, empirical evaluation remains limited. We seek to extend preliminary results in narrative-centered learning by reporting on a controlled experiment assessing learning outcomes between several versions of a NLE and a more traditional, didactic format.

3 Crystal Island

CRYSTAL ISLAND is a narrative-centered learning environment built on Valve Software's Source™ engine, the 3D game platform for Half-Life 2. CRYSTAL ISLAND features a science mystery set on a recently discovered volcanic island. The curriculum underlying CRYSTAL ISLAND's science mystery is derived from the North Carolina state standard course of study for eighth-grade microbiology. Students play the role of the protagonist, Alyx, who is attempting to discover the identity and source of an unidentified infectious disease plaguing a newly established research station. The story opens by introducing the student to the island and members of the research team for which the protagonist's father serves as the lead scientist. Several of the team's members have fallen gravely ill, including Alyx's father. Tensions have run high on the island, and one of the team members suddenly accuses another of having poisoned the other researchers. It is the student's task to discover the outbreak's cause and source, and either acquit or incriminate the accused team member.

CRYSTAL ISLAND's setting includes a beach area with docks, a large outdoor field laboratory, underground caves, and a research camp. Throughout the mystery, the student is free to explore the world and interact with other characters while forming questions, generating hypotheses, collecting data, and testing hypotheses. The student can pick up and manipulate objects, take notes, view posters, operate lab equipment, and talk with non-player characters to gather clues about the source of the disease. During the course of solving the mystery, the student is minimally guided through a five problem curriculum. The first two problems focus on pathogens, including viruses, bacteria, fungi, and parasites. The student gathers information by interacting with in-game pathogen "experts" and viewing books and posters in the environment. In the third problem, the student is asked to compare and contrast her knowledge of four types of pathogens. In the fourth problem, the student is guided through an inquiry-based hypothesis-test-and-retest problem. In this problem she must complete a "fact sheet" with information pertaining to the disease inflicting members of the CRYSTAL ISLAND research team. Once the "fact sheet" is completed and verified by the camp nurse, the student completes the final problem concerning the treatment of viruses, bacteria, fungi, and parasites, and selects the appropriate treatment plan for sickened CRYSTAL ISLAND researchers. The story and curriculum are interwoven throughout the interactive experience.

4 Method

4.1 Participants

There were 88 female and 91 male participants varying in age and race. Approximately 2% of the participants were American Indian or Alaska Native, 5% were Asian, 29% were Black or African American, 58% were Caucasian, 6% were Hispanic or Latino, and 6% were of other races. Participants were all eighth-grade students ranging in age from 12 to 15 (M = 13.27, SD = 0.51). The students had

recently completed the microbiology curriculum mandated by the North Carolina state standard course of study before receiving the instruments, tests, and interventions of this experiment.

4.2   Materials and Apparatus

The pre-experiment paper-and-pencil materials for each participant were completed one week prior to intervention. These materials consisted of a researcher generated CRYSTAL ISLAND curriculum test, demographic survey, Achievement Goals Questionnaire [4], Self-Efficacy for Self-Regulated Learning scale (SESRL) [3], Science Self-Efficacy scale, modified from [13], and Immersion Tendencies Questionnaire [21]. The CRYSTAL ISLAND curriculum test consists of 23 questions created by an interdisciplinary team of researchers and was approved for language and content by the students' eighth-grade science teachers. Elliot and McGregor's Achievement Goals Questionnaire is a validated instrument which measures four achievement goal constructs (mastery-approach, performance-approach, mastery-avoidance, and performance-avoidance goals) [4]. Bandura's Self-Efficacy for Self-Regulated Learning scale [3] consists of 11 items rated by participants on a 7-point Likert scale. Witmer and Singer developed and validated the Immersive Tendencies Questionnaire (ITQ) to measure individual predispositions towards presence experiences [21]. The ITQ consists of three subscales: activity involvement tendency, activity focus tendency, and video game playing tendency. Participants indicate their level of tendency for each item on a 7-point Likert scale. Witmer and Singer found individual tendencies, as recorded by the ITQ, to be predictive of presence (discussed in Section 6.2) [21].

Post-experiment materials were completed immediately following intervention. These materials consisted of the same CRYSTAL ISLAND curriculum test, Achievement Goals Questionnaire [4], Science Self-Efficacy scale, an interest scale [19], and the Presence Questionnaire [21]. The interest scale was adapted from those used by Schraw to capture differences across groups and to examine within-subject relationships with learning outcomes [19]. Participants' presence experience was captured with the Presence Questionnaire (PQ) developed and validated by Witmer and Singer [21]. The PQ contains several subscales including involvement/control, naturalism of experience and quality of the interface scales.

5   Design and Procedure

5.1   Design

The experiment randomly assigned the entire eighth grade population of Centennial Campus Middle School in Raleigh, North Carolina to four groups: holdout, CRYSTAL ISLAND narrative condition, CRYSTAL ISLAND minimal-narrative condition, or PowerPoint condition (see Table 1 for condition breakdown). Participants in the holdout condition did not receive an intervention and served as the control group for

this experiment and planned longitudinal studies. In the remaining three conditions, students received an intervention consisting of the CRYSTAL ISLAND microbiology curriculum delivered in one of three formats. The CRYSTAL ISLAND narrative condition supplemented the curriculum with the full CRYSTAL ISLAND narrative, including a poisoning scenario, character back-stories, and rich character personalities. The CRYSTAL ISLAND minimal-narrative condition supplemented the curriculum with the minimal story required to support the curriculum. In this condition, the story strictly consisted of research members falling ill and the request for the student to identify the mysterious illness. The minimal-narrative condition did not include the poisoning storyline, character back-stories, or explicit character personality. The PowerPoint condition consisted of a narrated PowerPoint presentation of the same curriculum that was used in CRYSTAL ISLAND. Much of the text and images of the slides actually appear in CRYSTAL ISLAND in the form of books, posters, and character dialogue. The PowerPoint condition did not contain a corresponding storyline.

Table 1. Subject population by condition.

Condition n = 179			
Holdout n = 29	Narrative n = 60	Min-narrative n = 56	PowerPoint n = 33
Female n = 18 Male n = 11	Female n = 30 Male n = 30	Female n = 24 Male n = 32	Female n = 16 Male n = 17

5.2   Participant Procedure

Participants entered the experiment room having completed the pre-test and instrumentation one week prior to the intervention. Participants were first instructed to review CRYSTAL ISLAND instruction materials. These materials consisted of the CRYSTAL ISLAND back-story and task description, a character handout, a map of the island, and a control sheet. Participants were then further directed on the controls via a presentation explaining each control in detail.
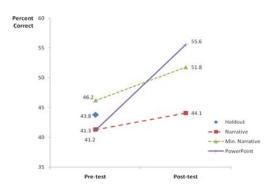
Participants in the three intervention conditions (narrative, minimal-narrative, and PowerPoint) were given 50 minutes to work on solving the mystery. Solving the mystery consisted of completing a number of goals including learning about pathogens, viruses, bacteria, fungi, and parasites, compiling the symptoms of the researchers who had fallen ill, recording features of hypothesized diseases causing the CRYSTAL ISLAND illness, testing a variety of possible sources, and reporting the solution (cause and source) to the camp nurse to design a treatment plan.

Immediately after solving the science mystery of CRYSTAL ISLAND, or 50 minutes of interaction, participants completed the post-experiment questionnaires. First to be completed was the CRYSTAL ISLAND curriculum test, followed by the remaining post-experiment questionnaires described above. Completion of post-experiment materials took no longer than 35 minutes for participants. In total, sessions lasted 90 minutes.

6   Results

6.1   Learning Outcomes

Investigating learning in CRYSTAL ISLAND as measured by the difference of post-test and pre-test scores, we find that, overall, students exhibited learning gains (learning gain, $M = 0.07$, $SD = 0.14$). On average, students answered 1.6 ($SD = 3.3$) more questions correctly on the post-test than on the pre-test. Matched pairs t tests (comparing post-test to pre-test scores) indicate that these learning gains are significant overall, $t(149) = 5.51$, $p < 0.0001$, and significant (although weakly significant in the narrative condition) within each condition (narrative condition: $t(58) = 1.43$, $p = 0.07$, minimal-narrative condition: $t(55) = 2.97$, $p < 0.005$, and the PowerPoint condition: $t(34) = 5.74$, $p < 0.0001$). Further, the learning gains in each condition were significantly different, $F(2, 149) = 10.38$, $p < 0.0001$. There was no significant difference among pre-test scores between conditions, $F(4, 179) = 0.94$, $p = 0.44$. The largest learning gains occurred in the PowerPoint condition ($M = 0.15$, $SD = 0.15$), followed by learning gains in the minimal-narrative condition ($M = 0.06$, $SD = 0.14$), and the lowest learning gains in the narrative condition ($M = 0.02$, $SD = 0.11$). Students in the hold out condition did not take a post-test, and therefore no learning gain results are available for that condition. The CRYSTAL ISLAND curriculum test consisted of 23 items leading to a learning gain of 0.043, which equates to getting one additional question correct in the post-test compared to the pre-test. Thus, on average, students in the PowerPoint condition answered 3.5 more questions correctly ($SD = 3.6$) on the post-test, with participants in the minimal-narrative and narrative conditions answering 1.3 ($SD = 3.2$) and 0.5 ($SD = 2.7$) more questions correctly, respectively. Learning gains are depicted in Figure 1. If we consider only the students who completed the CRYSTAL ISLAND mystery in the narrative condition, we find no significant difference between post-test scores with the PowerPoint condition, $F(1, 48) = 0.32$, $p = 0.58$. However, the learning gains in the PowerPoint condition were somewhat significantly better than the students who finished the CRYSTAL ISLAND narrative experience, $F(1, 48) = 4.09$, $p = 0.05$.

Fig. 1. Learning gains (pre- to post-test) by condition.

Interestingly, there was an effect of gender on learning in CRYSTAL ISLAND. When we consider only the problems on the CRYSTAL ISLAND curriculum test for which students were exposed to (not all students solved the CRYSTAL ISLAND mystery and completed all problem-solving activities), we find gender played a significant role, $F(1, 114) = 4.44$, $p = 0.037$. In

CRYSTAL ISLAND, on average, male students got an additional 1.3 problems correct (SD = 2.4) on post-tests compared to pre-tests, while female students got an additional 0.4 problems correct (SD = 1.7).

6.2 Presence Outcomes

Presence contributes to the goal of transparency in technology-mediated interactions. 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 [5, 17]. Presence has been alternatively defined as "the perceptual illusion of nonmediation" [9], as well as "the subjective experience of being in one place or environment, even when one is physically situated in another" [21]. Witmer and Singer further distinguish presence from involvement. *Involvement* refers to the degree of attention and meaning devoted to some set of stimuli [21]. Here we report on students' reported sense of presence while interacting in the CRYSTAL ISLAND storyworld (narrative and minimal-narrative conditions only).

Narrative had a significant effect on student presence, $F(1, 115) = 4.23$, $p = 0.04$. Higher presence was reported in the narrative condition ($M = 147.35$, $SD = 30.6$) compared to the minimal-narrative condition ($M = 136.5$, $SD = 25.8$). Gender was also found to have a weakly significant effect on presence, $F(1, 115) = 2.87$, p = 0.09, with females reporting higher levels of presence ($M = 146.9$, $SD = 26.1$) than males ($M = 137.9$, $SD = 30.5$). Students reporting high-levels of interest (as gauged by the interest scale modified from [19]) reported higher levels of presence than students with low-levels of interest. There was a significant correlation of interest with student presence, $r(114) = 0.36$, $p = 0.0001$, and several of the PQ's subscales, including: involvement/control ($r(114) = 0.42$, $p < 0.0001$), naturalism of experience ($r(114) = 0.27$, $p = 0.003$), and resolution ($r(114) = 0.29$, $p = 0.002$). Self-efficacy and presence also had a significant interaction. Students with high science efficacy reported higher levels of presence than less efficacious students, $r(114) = 0.35$, $p = 0.0001$. Likewise, students reporting greater levels of involvement and control (a PQ subscale) also reported higher science efficacy, $r(114) = .28$, $p = 0.002$.

Student goal orientation was found to affect presence as well. In particular, there was a significant effect of mastery approach on presence in both CRYSTAL ISLAND conditions, $F(1, 114) = 8.65$, $p = 0.004$, and performance avoidance on presence, $F(1, 114) = 4.59$, $p = 0.034$. Mastery oriented students reported greater levels of presence than performance-oriented students. Students who sought to avoid negative performance outcomes also reported higher levels of presence than students who did not seek to avoid negative performance outcomes.

7 Discussion

The experiment found that students who interacted with the CRYSTAL ISLAND environment achieved significant learning gains. While pre- to post-test performance differences were greatest in the PowerPoint condition, the findings support the hypothesis that students received clear motivational benefits from interacting with

CRYSTAL ISLAND. Further, student levels of presence had significant relationships with factors relevant to learning and motivation, including self-efficacy, interest, involvement/control, and goal orientation. While learning gains were higher in the minimal-narrative condition, students reported higher levels of presence in the narrative condition, carrying promising implications for motivation.

Drawing upon the experiment's learning gain results, it is possible that the narrative condition's additional story content overloaded cognition, enabling students to learn more without the supplemental storyline on proximal assessment. An important direction for future work is conducting longitudinal studies to determine the the long-term effects of narrative on learning and inform scaffolding strategies for reducing cognitive load.

The study found significant effects of presence, with higher reported presence in the narrative condition. With the benefits of efficacious learners having been widely demonstrated [2,22], it is important to note that higher presence levels also lead to higher levels of reported self-efficacy. If further study can identify the narrative factors that contribute to motivation and efficacy, we can enhance the ability of NLEs to support student problem solving, increase student effort, persistence, and resilience when confronted with failure, and raise the levels of success students are likely to achieve [2, 18, 22].

When considering the Involvement/Control subscale of the Presence Questionnaire [21], the findings indicated that high levels of Involvement/Control are correlated with higher reports of self-efficacy. Perception of control is known to have motivational benefits [10]. As a factor contributing to presence, involvement/control suggests probable relationships between presence and motivation. The findings of this study highlight the potential for positive connections between narrative and motivation that deserve further investigation. Further exploration of these relationships will contribute to a deepened understanding of the narrative factors that relate story content, presence, learning, motivation, and self-efficacy, as well as our ability to regulate these factors in an effort to support pedagogical objectives.

The study also found an effect of student goal orientation on perceptions of presence among the middle school participants. The gaming environment, on which CRYSTAL ISLAND is built, may have had an effect on performance-oriented students, encouraging them to attempt to solve the mystery quickly. Meanwhile, it seems that mastery oriented students, who tend to measure accomplishments by learning successes, reported a greater perception of presence. It is probable that mastery oriented students were more likely to take their time throughout their interactions, focusing their attention on the content of learning environment so that their presence experience was heightened.

8   Limitations

The experiment was designed to control for time on task, allowing 50 minutes for the intervention. As a result of this constraint and the amount of content in CRYSTAL ISLAND, only 49 of the 116 CRYSTAL ISLAND participants finished or were working on the final problem at the end of the 50 minute session. An alternative design, which

will be adopted in future work, would consider controlling for task completion. Another limitation is that this study, at the time of writing, does not include a longitudal test to assess the hypothesized benefits of narrative.

9   Conclusion

Narrative is receiving increasing attention in the ITS community as a medium for contextualizing learning in meaningful ways while creating rich, engaging experiences for learners. To date, there has been little empirical work supporting the use of narrative in interactive learning environment. In a controlled experiment with an "early-generation" NLE, it was found that students do in fact exhibit learning gains, that those gains are less than those in produced by traditional instructional approaches, but that the motivational benefits of narrative-centered learning, particularly with regard to self-efficacy, presence, interest, and perception of control, are substantial.

  The results highlight two important directions for future work. First, the contextual benefits of narrative may be more pronounced in a longitudinal evaluation of learning rather than in the assessment administered immediately following intervention as in the study reported here. Second, it is important to begin exploring the educational role of the myriad components of narrative in learning episodes, such as plot coherence, drama, and character identification, and their impact on problem solving, learning outcomes, and engagement [16].

Acknowledgments

References

1.   Aylett, R., Louchart, S., Dias, J., Paiva, A., and Vala, M. FearNot! - An Experiment in Emergent Narrative. In *Proceedings of IVA 2005*, pp. 305-316, Kos, Greece: Springer (2005)
2.   Bandura, A. *Self-efficacy: The exercise of control*. New York: Freeman (1997)
3.   Bandura, A. Guide for constructing self-efficacy scales. In F. Pajares and T. Urdan (eds.) *Self-Efficacy Beliefs of Adolescents,* pp.307-337. Greenwich, Connecticut: Information Age Publishing (2006)
4.   Elliot, A., and McGregor, H. A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology*, 80, 3, 501-519 (2001)

5. Insko, B. Measuring presence: Subjective, behavioral and physiological methods. In G. Riva, F. Davide, and W. IJsselsteijn (Eds.), *Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments,* pp.109-119. Amsterdam: Ios Press (2003)

6. Johnson, W. L, and Beal, C. Iterative Evaluation of an Intelligent Game for Language Learning. In *Proceedings of AIED 2005*, Amsterdam: IOS Press (2005)

7. Johnson, W. L. Serious Use of a Serious Game for Language Learning. In *Proceedings of AIED 2007*, Marina del Rey, CA: IOS Press (2007)

8. Linnenbrink, E. A., and Pintrich, P. R. Multiple goals, multiple contexts: The dynamic interplay between personal goals and contextual goal stresses. In S. Volet and S. Jarvela (Eds.) *Motivation in Learning Contexts: Theoretical Advances and Methodological Implications,* pp. 251-269, New York: Elsevier (2001)

9. Lombard, M., and Ditton, T. At the heart of it all: The concept of presence. *Journal of Computer Mediated Communication*, 3, 2 (1997)

10. Malone, T., and Lepper, M. Making learning fun: A taxonomy of intrinsic motivations for learning. In Snow, R., and Farr, M. (Eds.), *Aptitude, Learning, and Instruction: III. Conative and Affective Process Analyses,* pp. 223-253, Hillsdale, NJ: Erlbaum (1987)

11. Marsella S., Johnson, W.L. and LaBore, C. Interactive pedagogical drama for health interventions. In *Proceedings of the 11th International Conference on Artificial Intelligence in Education*, Australia (2003)

12. Mott, B., and Lester, J. Narrative-centered tutorial planning for inquiry-based learning environments. In *Proceedings of the 8th International Conference on Intelligent Tutoring Systems*, 675-684, Jhongli, Taiwan (2006)

13. Nietfeld, J. L., Cao, L., and Osborne, J. W. The effect of distributed monitoring exercises and feedback on performance and monitoring accuracy. *Metacognition and Learning*, 1, 2, 159-179 (2006)

14. Perry, N. E. Young children's self-regulated learning and the contexts that support it. *Journal of Educational Psychology,* 90, 715-729 (1998)

15. Riedl, M. and Stern, A. Believable agents and intelligent story adaptation for interactive storytelling. In *Proceedings of the 3rd International Conference on Technologies for Interactive Digital Storytelling and Entertainment*, Darmstadt, Germany (2006)

16. Rowe, J., McQuiggan, S., and Lester, J. Narrative presence in intelligent learning environments. In *Working Notes of the 2007 AAAI Fall Symposium on Intelligent Narrative Technologies*, Washington DC (2007).

17. Schubert, T., Friedmann, F., and Regenbrecht, H. Embodied presence in virtual environments. In Paton, R., and Neilson, I. (Eds.), *Visual Representations and Interpretations*, pp. 269-278, London: Springer (1999)

18. Schunk, D. and Pajares, F. The development of academic self-efficacy. In Wigfield, A. and Eccles, J. (Eds.), *Development of Achievement Motivation,* pp. 15-31, San Diego, CA: Academic Press (2002)

19. Schraw, G. Situational interest in literary text. *Contemporary Educational Psychology*, 22, 436–456 (1997)

20. Wells, C. *The Meaning Makers: Children Learning Language and Using Language to Learn*. Portsmouth, NH: Heinnemann (1986)

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

22. Zimmerman, B. Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25, 82-91 (2000)