Predicting Student Self-Regulation Strategies in Game-Based Learning Environments

Jennifer Sabourin, Lucy Shores, Bradford Mott, and James Lester

North Carolina State University, Raleigh, North Carolina, USA
{jlrobiso, lrshores, bwmott, lester}@ncsu.edu

Abstract. Self-regulated learning behaviors such as goal setting and monitoring have been found to be key to students' success in a broad range of online learning environments. Consequently, understanding students' self-regulated learning behavior has been the subject of increasing interest in the intelligent tutoring systems community. Unfortunately, monitoring these behaviors in real-time has proven challenging. This paper presents an initial investigation of self-regulated learning in a game-based learning environment. Evidence of goal setting and monitoring behaviors is examined through students' text-based responses to update their 'status' in an in-game social network. Students are then classified into SRL-use categories that can later be predicted using machine learning techniques. This paper describes the methodology used to classify students and discusses initial analyses demonstrating the different learning and gameplay behaviors across students in different SRL-use categories. Finally, machine learning models capable of predicting these categories early into the student's interaction are presented. These models can be leveraged in future systems to provide adaptive scaffolding of self-regulation behaviors.

Keywords: Self-regulated learning, machine learning, early prediction

1 Introduction

Understanding and facilitating students' self-regulated learning behaviors has been the subject of increasing attention in recent years. This line of investigation is fueled by evidence suggesting the strong role that self-regulatory behaviors play in a student's overall academic success [1]. Self-regulated learning (SRL) can be described as "the process by which students activate and sustain cognitions, behaviors, and affects that are systematically directed toward the attainment of goals" [2]. Unfortunately, students can demonstrate a wide range of fluency in their SRL behaviors [3] with some students lagging behind their peers in their ability to appropriately set and monitor learning goals.

For this reason, the ability to identify and support students' SRL strategies has been the focus of much work in the intelligent tutoring systems community [4,5,6]. Such work has focused primarily on examining SRL in highly structured problem-solving and learning environments. However, understanding and scaffolding students'

adfa, p. 1, 2011.
© Springer-Verlag Berlin Heidelberg 2011

SRL behaviors is especially important in open-ended learning environments where goals may be less clear and students do not necessarily have a clear indicator of their progress. In order to be successful in this type of learning environment, students must actively identify and select their own goals and evaluate their progress accordingly. Unfortunately, students do not consistently demonstrate sufficient self-regulatory behaviors during interactions with these environments, which may reduce the educational potential of these systems [7,8]. Consequently, further investigation of the role of SRL in open-ended learning environments is crucial for understanding how these environments can be used as effective learning tools.

This work describes a preliminary investigation of self-regulatory behaviors of students in a game-based science mystery, CRYSTAL ISLAND. During interactions with the CRYSTAL ISLAND environment, students were prompted to report on their mood and status in a way that is similar to many social networking tools available today. Though students were not explicitly asked about their goals or progress, many students included this information in their short, typed status statements. This data is used to classify students into low, medium, and high self-regulated learning behavior classes. Based on these classifications we investigate differences in student learning and in-game behaviors in order to identify the role of SRL in CRYSTAL ISLAND. Machine learning models are then trained that are capable of accurately predicting students' SRL-use categories early into their interaction with the environment, offering the possibility for timely intervention. The implications of these results and areas of future work are then discussed.

2 Related Work

Self-regulated learning (SRL) is a term used to describe the behaviors of students who actively control their learning goals and outcomes [9]. Among other things, SRL involves students actively setting goals and making conscious choices to measure and evaluate their progress towards them. Self-regulated learners deliberately reflect on their knowledge and learning strategies and make adjustments based on past success and failure. While it seems all students apply self-regulatory behaviors during learning, the degree of competency is unfortunately broad, even among students of the same age [3]. Additionally, there is evidence that individuals who are better able to regulate their learning in intentional and reflective ways are more likely to achieve academic success [1]. To mediate these differences, intervention research focused on process goals and feedback has been conducted in traditional classrooms and has yielded positive results [9,10,11].

Beyond the traditional classroom, identifying and scaffolding SRL strategies has been a focus of much work in the intelligent tutoring systems community as well. For example, in MetaTutor, a hypermedia environment for learning biology, think-aloud protocols have been used to examine which strategies students use, while analysis of students' navigation through the hypermedia environment helps to identify profiles of self-regulated learners [6]. Similarly, researchers have identified patterns of behavior in the Betty's Brain system that are indicative of low and high levels of self-regulation [5]. Prompting students to use SRL strategies when these patterns of behavior occur has shown promise in improving student learning. Conati *et al*. have examined the benefits of prompting students to self-explain when learning physics content in a computer-based learning environment [4].

While previous work has focused primarily on examining SRL in highly structured problem-solving and learning environments, there has also been work on identifying SRL behaviors in open-ended exploratory environments. For example, work by Shores *et al.* has examined early prediction of students' cognitive tool use in order to inform possible interventions and scaffolding [12]. Understanding and scaffolding student's SRL behaviors is especially important in open-ended learning environments where goals may be less clear and students do not necessarily have a clear indicator of their progress [13]. In order to be successful in this type of learning environment, students must actively identify and select their own goals and evaluate their progress accordingly. While the nature of the learning task may have implicit overarching goals such as 'completing the task' or 'learning a lot,' it is important for students to set more specific, concrete and measurable goals [14].

This work focuses on examining SRL within the context of narrative-centered learning. *Narrative-centered learning environments* are a class of serious games that tightly couple educational content and problem solving with interactive story scenarios. By contextualizing learning within narrative settings, narrative-centered learning environments tap into students' innate facilities for crafting and understanding stories [15]. Narrative-centered learning environments have been developed that teach negotiation skills [16] and foreign languages [17] through conversational interactions with virtual characters. Scientific inquiry has been realized in interactive mysteries where students play the roles of detectives [18,19]. While these environments are capable of providing rich, engaging experiences [18], they should not overload students by providing too many possible paths for learning [7]. Appropriate goal-setting is necessary to succeed in these learning environments, making the ability to recognize and support students' SRL strategies especially critical.

3 Method

An investigation of students' SRL behaviors was conducted with CRYSTAL ISLAND, a game-based learning environment being developed for the domain of microbiology that follows the standard course of study for eighth grade science in North Carolina. CRYSTAL ISLAND features a science mystery set on a recently discovered volcanic island. Students play the role of the protagonist, Alex, who is attempting to discover the identity and source of an unidentified disease plaguing a newly established research station. The story opens by introducing the student to the island and the members of the research team for which her father serves as the lead scientist. As members of the research team fall ill, it is her task to discover the cause and the specific source of the outbreak. Typical game play involves navigating the island, manipulating objects, taking notes, viewing posters, operating lab equipment, and talking with non-player characters to gather clues about the disease's source. To progress through the

Table 1. SRL Tagging Scheme

SRL Category	Description	Examples
Specific reflection	Student evaluates progress towards a specific goal or area of knowledge	"I am trying to find the food or drink that caused these people to get sick." "Well...the influenza is looking more and more right. I think I'll try testing for mutagens or pathogens – [I] ruled out carcinogens"
General reflection	Student evaluates progress or knowledge but without referencing a particular goal	"I think I'm getting it" "I don't know what to do"
Non-reflective	Student describes what they are doing or lists a fact without providing an evaluation	"testing food" "in the lab"
Unrelated	Any statement which did not fall into the above three categories is considered unrelated, including non-word or unidentifiable statements	"having fun" "arghhh!"

mystery, a student must explore the world and interact with other characters while forming questions, generating hypotheses, collecting data, and testing hypotheses.

A study with 296 eighth grade students was conducted. Participants interacted with CRYSTAL ISLAND in their school classroom, although the study was not directly integrated into their regular classroom activities. Pre-study materials were completed during the week prior to interacting with CRYSTAL ISLAND. The pre-study materials included a demographic survey, researcher-generated CRYSTAL ISLAND curriculum test, and several personality questionnaires including *personality* [20] and *goal orientation* [21]. Students were allowed approximately 55 minutes to attempt to solve the mystery. Immediately after solving the mystery, or after 55 minutes of interaction, students moved to a different room in order to complete several post-study questionnaires including the curriculum post-test.

Students' affect data were collected during the learning interactions through self-report prompts. Students were prompted every seven minutes to self-report their current mood and status through an in-game smartphone device. Students selected one emotion from a set of seven options, which consisted of the following: *anxious*, *bored*, *confused*, *curious*, *excited*, *focused*, and *frustrated*. After selecting an emotion, students were instructed to briefly type a few words about their current status in the game, similarly to how they might update their status in an online social network. These status reports were later tagged for SRL evidence use using the following four ranked classifications: (1) *specific reflection*, (2) *general reflection*, (3) *non-reflective statement*, or (4) *unrelated* (Table 1). This ranking was motivated by the observation that setting and reflecting upon goals is a hallmark of self-regulatory behavior and that specific goals are more beneficial than those that are more general [14]. Students were then given an overall SRL score based on the average score of their statements. An even tertiary split was then used to assign the students to a Low, Medium, and High SRL category.

4 Results

Data was collected from 296 eighth grade students from a rural North Carolina middle school. After removing instances with incomplete data or logging errors, there were 260 students remaining. Among the remaining students, there were 129 male and 131 female participants varying in age and ethnicity. A total of 1836 statements were collected, resulting in an average of 7.2 statements per student. All statements were tagged by one member of the research team with a second member of the research team tagging a randomly selected subset (10%) of the statements to assess the validity of the protocol. Inter-rater reliability was measured at $\kappa = 0.77$, which is an acceptable level of agreement. General reflective statements were the most common (37.2%), followed by unrelated (35.6%), specific reflections (18.3%) and finally non-reflective statements (9.0%).

4.1 Analyzing Self-Regulation Behaviors

The first objective of this investigation was to explore differences in student learning based on self-regulatory tendencies. Student learning, as measured by normalized learning gains from the pre-test to post-test, was compared for the three SRL groups. An ANOVA indicated a difference in learning gains between the groups ($F_{(2, 257)} = 4.6$, $p < 0.01$). Tukey post-hoc comparisons indicated that both High and Medium SRL students experienced significantly better learning gains than Low SRL students at the $\alpha = 0.05$ level. Analyses also indicated that there were significant differences on pre-test scores between groups ($F_{(2, 257)} = 5.07$, $p < 0.01$) suggesting that students with high SRL tendencies may be better students or perhaps their increased prior knowledge helped them to identify and evaluate their goals more efficiently. Figure 1 shows the pre- and post- test scores across groups, highlighting both the differences in pre-knowledge and learning during interaction with CRYSTAL ISLAND.

The next set of analyses was conducted to investigate differences in student behavior based on their SRL tendencies. A chi-squared analysis indicated that the percentage of students who solved the mystery did not differ significantly based on SRL

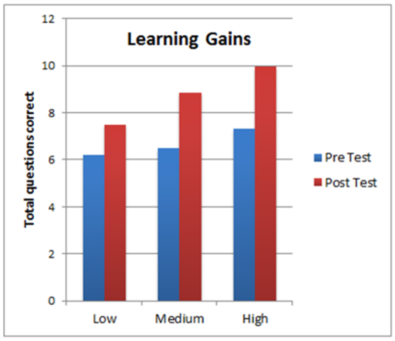

Fig. 1. Learning gains by SRL group

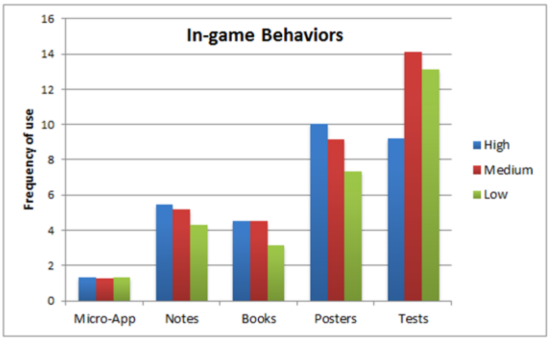

Fig 2. In-game behaviors by SRL group

group ($\chi 2$ (2, N=260) = 4.72, p = 0.094). Additionally, an ANOVA indicated there was no significant difference in the number of goals completed during the interaction.

While a significant difference in students' abilities to solve the mystery was not found, there were differences in the in-game resources that students used. Resources expected to be most beneficial to learning and self-regulation included a *microbiology app* on the students' in-game smartphone which provides a wealth of microbiology information, *books* and *posters* that are scattered around the island with additional information, a notebook where students can record their own *notes*, and finally a *testing* machine where students formulate hypotheses and run the relevant tests. ANOVAs for student use of each of these features indicated a significant difference in student use of *posters* ($F_{(2, 257)} = 5.28$, $p < 0.01$), and *tests* ($F_{(2, 257)} = 5.59$, $p < 0.01$). While the differences in the use of other devices were not significant, interesting trends emerged (Figure 2). High SRL students appear to make more use of the curricular resources in the game such as books and posters and also take more notes than the lower SRL students. Interestingly, High SRL students run significantly fewer tests than Medium or Low SRL students (as indicated by Tukey post-hoc comparisons). Abundant use of the testing device is often indicative of students gaming the system or failing to form good hypotheses in advance. This finding suggests that High SRL students may be more carefully selecting which tests to run and are perhaps obtaining positive test results earlier than Medium and Low SRL students.

4.2 Predicting Self-Regulation Behaviors

These results highlight several important factors relating to self-regulation. First, the post-interaction method of classifying students into Low, Medium, and High SRL categories appears to yield meaningful groupings of students. Second, these classifications have significant implications for student learning. Students in the High SRL group have a higher level of initial knowledge than Low SRL students and through interactions with CRYSTAL ISLAND, increase this gap in knowledge. This highlights the importance of identifying the Low SRL students so they can receive supplementary guidance to help bridge this gap. Finally, the results indicate that High SRL students utilize the environment's curricular features differently and likely more effectively than Low SRL students. This finding suggests that scaffolding to direct Low SRL students towards more effective use of these resources could be an appropriate mechanism for bridging the learning gap.

However, in order to make use of these findings, Low SRL students must be identified early into the interaction so they can be provided with the necessary scaffolding. The current procedure for identifying these students is performed manually after the interaction has been completed, which does not allow for early interventions. It is also desirable to only provide additional scaffolding to the Low SRL students since the other students appear to be effectively using the environment already and may potentially be harmed by additional interventions. For these reasons, the next goal of this research was to train machine-learning models to predict students' SRL-use categories early into their interaction with CRYSTAL ISLAND.

Table 2. Predictive models and evaluation metrics (for predictive accuracy, * and ** indicate a significant improvement over the prior prediction at p <.05 and .01, respectively)

Model	*Predictive Accuracy*				*Low-SRL Recall*			
	Initial	$Report_1$	$Report_2$	$Report_3$	Initial	$Report_1$	$Report_2$	$Report_3$
Naïve Bayes	44.2	43.5	46.1*	50.5*	0.47	0.28	0.54	0.52
Neural Network	42.3	43.8	46.5*	45.5	0.44	0.45	0.49	0.52
Log. Reg.	42.7	51.2**	47.7	54.5**	0.45	0.65	0.66	0.73
SVM	43.5	46.9*	45.7	51.4**	0.51	0.55	0.56	0.62
Decision Tree	42.7	46.2*	48.1*	57.2**	0.45	0.55	0.71	0.71

In order to predict students' SRL-use categories, a total of 49 features were used to train machine-learning models. Of these, 26 features represented personal data collected prior to the student's interaction with CRYSTAL ISLAND. This included demographic information, pre-test score, and scores on the personality, goal orientation, and emotion regulation questionnaires. The remaining 23 features represented a summary of students' interactions in the environments. This included information on how students used each of the curricular resources, how many in-game goals they had completed, as well as evidence of off-task behavior. Additionally, data from the students' self-reports were included, such as the most recent emotion report and the character count of their "status."

In order to examine early prediction of the students' SRL-use categories, these features were calculated at four different points in time resulting in four distinct datasets. The first of these (**Initial**) represented information available at the beginning of the student's interaction and consequently only contained the 26 personal attributes. Each of the remaining three datasets (**$\mathbf{Report_{1\text{-}3}}$**) contained data representing the student's progress at each of the first three emotion self-report instances. These datasets contained the same 26 personal attributes, but the values of the remaining 23 in-game attributes differentially reflected the student's progress up until that point. The first self-report occurred approximately 4 minutes into game play with the second and third reports occurring at 11 minutes and 18 minutes, respectively. The third report occurs after approximately one-third of the total time allotted for interaction has been completed, so it is still fairly early into the interaction time.

Each of these datasets was used to train a set of machine learning classifiers including: Naïve Bayes, Decision Tree, Support Vector Machine, Logistic Regression, and Neural Network. These models were trained and evaluated using 10-fold cross-validation with the WEKA machine learning toolkit [22]. The predictive accuracies of these models are shown in Table 2. All of the learned models were able to offer a predictive accuracy statistically significantly better than a most-frequent class baseline (at $p < 0.01$). Due to the fact that the classes were identified using an even tertiary split, the most frequent class (Medium) model has a predictive accuracy of 33.5%. Additionally, most models demonstrated gains in predictive accuracy further into the interaction.

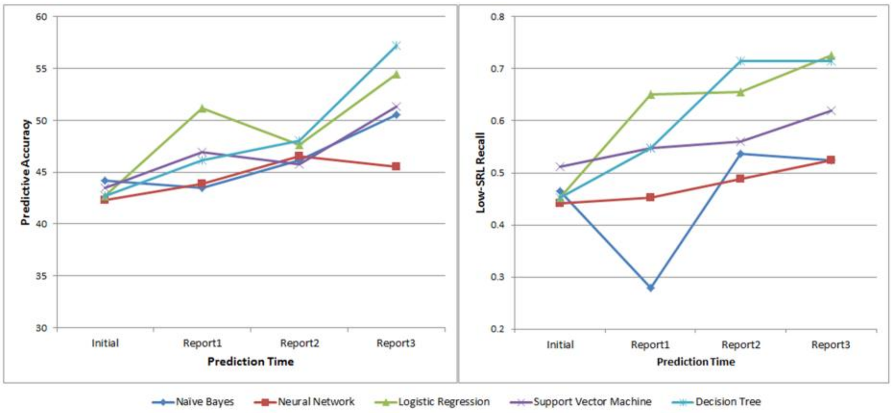

Fig. 3. Predictive accuracy and Low-SRL recall improvements across time

Of the models attempting to predict SRL class before any interaction with the environment, the model with the best performance is the Naïve Bayes model (44.3%). However, there are no significant differences in predictive accuracy between any of the models trained on this dataset. Alternatively, of the models trained with the most data, the Decision Tree model achieves the highest predictive accuracy (57.2%), and is statistically significantly better than the other models trained on this dataset ($p < 0.05$). In general, it appears that the two models with the best overall performance are the Decision Tree and Logistic Regression models.

In addition to predictive accuracy, we are also particularly interested in the models' abilities to distinguish Low SRL students as these students would be the targets of additional support. For this reason, we compared the models' levels of recall for the Low SRL class (Figure 3). These results again demonstrate a steady growth in the ability to correctly recognize Low SRL students. Additionally, the Decision Tree and Logistic Regression models again distinguish themselves in their ability to outperform the remaining models. These results indicate that using either model, or perhaps a combination of both models, will offer promise in being able to identify and support Low SRL students early into their interaction with CRYSTAL ISLAND.

5 Discussion

This work presents an initial analysis of students' natural self-regulated learning activities in the narrative-centered learning environment, CRYSTAL ISLAND. Results indicate that undirected prompts have the potential to show students' use of goal setting and monitoring. Additionally, the findings suggest that self-regulated learners tend to make better use of in-game curricular resources and may be more deliberate in their actions. Though highly self-regulated learners were not more likely to solve the mystery, they did demonstrate significantly higher learning gains as a result of their interaction. These results point to the importance of being able to identify students with tendencies towards low self-regulation in order to provide appropriate scaffolding. The machine learning models discussed in this paper show significant promise in being able to predict a student's SRL abilities early into their interaction with CRYSTAL ISLAND.

These findings point to several natural directions for future work. The most prominent of these is developing intervention mechanisms for aiding student self-regulation. Specifically, the results of this work point to the ways that in-game curricular resources can be used effectively. Low SRL students could receive additional support in their use of these resources. Alternatively, it may be that these students suffer in their abilities to recognize and set appropriate goals. This goal-setting behavior could be made more explicit using the game-based nature of the environment.

Understanding how to effectively incorporate these strategies into narrative-centered learning environments is an important area for future investigation. Drawing on ongoing empirical investigations of learning, problem solving, and engagement can support the exploration of a broad range of potential techniques for further enhancing student SRL skills. In particular, investigating individualized instruction strategies and designing SRL features for narrative environments that account for individual differences is an important next step in this line of investigation.

Acknowledgments. The authors wish to thank members of the IntelliMedia Group for their assistance, Omer Sturlovich and Pavel Turzo for use of their 3D model libraries, and Valve Software for access to the Source™ engine and SDK. This research was supported by the National Science Foundation under Grants REC-0632450, DRL-0822200, and IIS-0812291. This material is based upon work supported under a National Science Foundation Graduate Research Fellowship. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. B. J. Zimmerman, "Self-regulated learning and academic achievement: An overview," *Educational Psychologist*, vol. 25, pp. 3-17, 1990.
2. D. H. Schunk, "Attributions as Motivators of Self-Regulated Learning," in *Motivation and Self-Regulated Learning: Theory, Research, and Applications*, D. H. Schunk and B. J. Zimmerman, Eds. 2008, pp. 245-266.
3. D. Ellis and B. J. Zimmerman, "Enhancing self-monitoring during self-regulated learning of speech," H. J. Hartman, Ed. Dordrecht, The Netherlands: Kluwer, 2001, pp. 205-228.
4. C. Conati and K. VanLehn, "Towards Computer-Based Support of Meta-Cognitive Skills: a Computational Framework to Coach Self-Explanation," *International Journal of Artificial Intelligence in Education*, vol. 11, pp. 398-415, 2000.
5. G. Biswas, H. Jeong, R. Roscoe, and B. Sulcer, "Promoting Motivation and Self-Regulated Learning Skills through Social Interactions in Agent-Based Learning Environments," in *2009 AAAI Fall Symposium on Cognitive and Metacognitive Educational Systems*, 2009.
6. R. Azevedo, A. Johnson, A. Chauncey, and C. Burkett, "Self-Regulated Learning with MetaTutor: Advancing the Science of Learning with MetaCognitive Tools," in *New Sci-*

ence of Learning: Cognition, Computers and Collaboration in Education, M. Khine and I. Saleh, Eds. New York: Springer, 2010, pp. 225-248.
7. P. A. Kirschner, J. Sweller, and R. E. Clark, "Why Minimal Guidance during instruction does not work: An analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching," *Educational Psychologist*, vol. 41, pp. 75-86, 2006.
8. L. Alfieri, P. Brooks, N. Aldrich, and H. Tenenbaum, "Does Discovery-Based Instruction Enhance Learning," *Journal of Education Psychology*, vol. 103, no. 1, pp. 1-18, 2011.
9. D. H. Schunk and B. J. Zimmerman, "Self-regulation and learning," vol. 7, W. M. Reynolds and G. E. Miller, Eds. New York, NY: Wiley & Sons, 2003, pp. 59-78.
10. D. H. Schunk and C. W. Swartz, "Goals and progress feedback: Effects on self-efficacy and writing achievement," *Contemporary Educational Psychology*, vol. 18, no. 3, pp. 337-354, 1993.
11. D. H. Schunk and C. W. Swartz, "Writing strategy instruction with gifted students: Effects of goals and feedback on self-efficacy and skills," *Roeper Review*, vol. 15, no. 4, pp. 225-230, 1993.
12. L. R. Shores, J. P. Rowe, and J. C. Lester, "Early Prediction of Cognitive Tool Use in Narrative-Centered Learning Environments," in *Proceedings of the Fifteenth International Conference on Artificial Intelligence in Education*, 2011.
13. S. Land, "Cognitive requirements for learning with open-ended learning environments," *Educational Technology Research and Development*, vol. 48, no. 3, pp. 61-78, 2000.
14. B. Zimmerman, "Goal Setting: A Key Proactive Source o Academic Self-Regulation," in *Motivation and Self-Regulated Learning: Theory, Research, and Applications*, D. H. Schunk and B. J. Zimmerman, Eds. New York: Routledge, 2008, pp. 267-286.
15. J. S. Bruner, *Acts of Meaning*. Cambridge, MA: Harvard University Press, 1990.
16. J. Kim et al., "BiLAT: A game-based environment for practicing negotiation in a cultural context," *International Journal of Artificial Intelligence in Education*, vol. 19, no. 3, pp. 289-308, 2009.
17. W. L. Johnson, "Serious Use of a Serious Game for Language Learning," in *Proceedings of the Thirteenth International Conference on Artificial Intelligence in Education*, 2007, pp. 67-74.
18. J. P. Rowe, L. R. Shores, B. W. Mott, and J. C. Lester, "Integrating Learning, Problem Solving, and Engagement in Narrative-Centered Learning Environments," *International Journal of Artificial Intelligence in Education*, 2011.
19. D. J. Ketelhut, "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*, vol. 16, no. 1, pp. 99-111, 2007.
20. R. McCrae and P. Costa, *Personality in Adulthood: A Five-Factor Theory Perspective*, 2nd ed. New York: Guilford Press, 1993.
21. A. Elliot and H. A. McGregor, "A 2 x 2 achievement goal framework," *Journal of Personality and Social Psychology*, vol. 80, pp. 501-519, 2001.
22. M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutmann, and I. Witten, "The WEKA Data Mining Software: An Update," *SIGKKD Explorations*, vol. 11, no. 1, 2009.