

# Affect and Engagement in Game-Based Learning Environments

Jennifer L. Sabourin and James C. Lester

**Abstract**—The link between affect and student learning has been the subject of increasing attention in recent years. Affective states such as flow and curiosity tend to have positive correlations with learning while negative states such as boredom and frustration have the opposite effect. Student engagement and motivation have also been shown to be critical in improving learning gains with computer-based learning environments. Consequently, it is a design goal of many computer-based learning environments to encourage positive affect and engagement while students are learning. Game-based learning environments offer significant potential for increasing student engagement and motivation. However, it is unclear how affect and engagement interact with learning in game-based learning environments. This work presents an in-depth analysis of how these phenomena occur in the game-based learning environment, CRYSTAL ISLAND. The findings demonstrate that game-based learning environments can simultaneously support learning and promote positive affect and engagement.

**Index Terms**—Games and infotainment, human factors

## 1 INTRODUCTION

INVESTIGATING the relationship between affect and learning has been the subject of increasing attention. This is perhaps due to early evidence that the relationship between emotion and cognitive activities is strong and complex. Affect has been shown to impact a variety of cognitive behaviors including decision making, information processing and moral reasoning [1]. Effective learning involves many such cognitive processes, such as attending to and incorporating new information, monitoring and evaluating one's own learning progress and creating plans for future behaviors to increase learning.

Research investigating the role of affect on learning has highlighted the importance of supporting positive affective states and aiding students in overcoming negative experiences. Specifically, it has been shown that positive affect, such as engaged concentration, joy and excitement, can lead to increased learning through better strategy selection [2], [3], increased persistence [4], and improved use of mental resources [5]. Alternatively, negative emotions, such as frustration, boredom and anger, are believed to lead to decreased motivation and effort [3], [6] and a desire to avoid the task [7].

Further work has examined how these phenomena occur in computer-based learning environments. These investigations have yielded evidence that students' emotional states can strongly impact how the student learns [8] and interacts with learning environments [9], [10]. Students with highly negative experiences such as frustration

and boredom are expected to persist in these negative states and may disengage from the learning task [10]. Bored students are particularly likely to engage in harmful behaviors such as gaming the system. There is also evidence that there may be some students who are better at regulating their affective experiences during learning [6]. For example, students who are focused on learning rather than objective measures of performance are more likely to recover from setbacks and states of confusion [11]. These findings again suggest that encouraging positive affect and student engagement is critical to learning and motivating students to pursue future learning tasks.

One approach to improving engagement and interest with learning environments is incorporating game features [12], [13], [14]. Digital games offer a multitude of mechanisms for motivating players and keeping them engaged both immediately and across multiple interactions [15], [16]. Additionally, many of these features naturally support learning goals. For example, reward structures and adaptive difficulty levels can lead to increased motivation in both game and learning environments [16], [17], [18].

However, while these systems have been the subject of increasing attention, the dynamics of affect, engagement and learning in game-based learning environments is a key open question in affective computing. This paper explores these issues in the game-based learning environment, CRYSTAL ISLAND, and summarizes empirical work that has been conducted along these lines. This paper begins by discussing relevant work in examining the relationship between affect, engagement and learning, as well as the role game-based learning has played in recent years. We then describe CRYSTAL ISLAND and a large corpus collection that was designed to gather data regarding student affect and engagement in the CRYSTAL ISLAND environment. We next discuss the roles of affect, engagement and learning in CRYSTAL ISLAND and describe how the game features support these processes. We then describe efforts to model these phenomena and the

- The authors are with the Department of Computer Science, North Carolina State University, Raleigh, NC 27695. E-mail: {jlrobiso, lester}@ncsu.edu.

Manuscript received 31 Aug. 2012; revised 5 Oct. 2013; accepted 11 Oct. 2013; date of publication 22 Oct. 2013; date of current version 9 May 2014.

Recommended for acceptance by G.N. Yannakakis, K. Isbister, A. Paiva, and K. Karpouzis.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.

Digital Object Identifier no. 10.1109/T-AFFC.2013.27

insight these efforts have yielded. Finally we discuss the implications of this work as well as future directions.

## 2 BACKGROUND

One-to-one tutoring has long been considered the gold standard of effective instruction by demonstrating significant improvements in student learning over the typical classroom setting [19]. This benefit is suspected to be due to high levels of interactivity as well as highly individualized attention and feedback. While it is not feasible for each student to have his or her own tutor, the intelligent tutoring systems community has worked to bridge this gap by endowing computers with the same interactive and individualized tutoring capabilities [20], [21]. However, as with any educational tool, their ability to produce learning gains for students depends on how motivated the student is to use them, as well as how engaged and effective students are during the interactions.

### 2.1 Student Affect

While the precise cognitive and affective mechanisms underlying learning experiences are not yet well understood, there has been significant progress in attempting to identify the emotions that students are likely to experience and how these may affect the learning process. For instance Kort et al. present a model of learning emotions that can be represented as a cycle that occurs throughout the learning process [22]. Other studies have investigated how emotional experiences transpire in computational environments. Both D'Mello et al. and Baker et al. have shown that students are most likely to remain in the same state through time and that certain emotional transitions are more likely than others [23], [24]. Their findings indicate that some negative states are particularly likely to persist and result in a "vicious cycle" of negative affect. Additionally, negative states such as frustration and especially boredom have been shown to have harmful effects on learning and lead students to disengage from the learning activities [9].

### 2.2 Student Engagement

Along these lines, there has been growing interest in how student motivation and engagement affects learning and problem solving. Of particular interest is answering the questions of how and why students disengage from educational software, as well as the cognitive impacts of disengagement [25], [26], [27]. Disengagement can take a variety of forms, including hint abuse [27], [28], off-task conversation [29] and gaming the system [26]. In general, students who abuse or disengage from an intelligent tutoring system learn less effectively than students who do not disengage [26], [30], [31]. Consequently, a growing body of research has investigated techniques for automatically detecting and preventing harmful learning behaviors such as gaming the system [27], [32], [33], [34].

Recent work investigating off-task behavior and student emotion has begun to raise questions about whether off-task behavior is universally unproductive for learning [35]. On the one hand, empirical findings suggest that off-task behavior is associated with boredom, which has been shown to be harmful for learning [9]. On the other hand,

recent findings have suggested that going off-task may alleviate negative affect, which could in turn benefit learning [35], [36]. A plausible explanation is that some students use off-task behavior as a coping strategy for negative learning emotions. These observations highlight the importance of further investigating how affect and engagement are related to behaviors that may be tangential to the learning task.

### 2.3 Game-Based Learning

Game-based learning has been proposed as an approach to encouraging positive affect, engagement and motivation in learning activities by utilizing game-like features and environments [15], [16], [18]. This work draws on empirical evidence that games are highly motivating and have natural ties with how people learn [16], [17], [18]. In recent years games have been used to teach a variety of subjects including scientific inquiry [14], [37], mathematics principles [13], negotiation skills [38], foreign languages [39], [40] policy argumentation [41] and critical reasoning [12].

Much work has examined whether these systems are effective at increasing student knowledge and skills, as well as fostering engagement and positive affect [42]. A study by Hallinen et al. indicated that incorporating game-like features into a learning system increased students' engagement though learning was the same between conditions [40]. On the other hand, a meta-analysis of serious games by Wouters et al. has shown strong learning benefits offered by game environments, with no motivational differences [42]. However, there is evidence that measures of motivation and engagement are strongly correlated with learning in game-based systems, as in traditional tutoring systems [14].

Recently, game-based learning has been criticized for having features that are superfluous to the learning task [43]. A major concern is that the game-play aspects that are designed to encourage interest and motivation may also introduce many distractions or "seductive details" that draw student attention away from the learning tasks [44], [45]. For example, students may become distracted by the characters and objects that are present in the world, or may spend time playing with aspects of the physics engine that underlies the gameplay. While this may be suspected to lead to more time spent off-task with less focus on learning, we hypothesize that this type of off-task behavior may be fundamentally different from off-task behavior in other learning environments. Underlying this difference is the idea that even when students are engaged in some off-task, in-game behaviors, they are still engaged with the environment. Unlike typical off-task behavior, which often involves conversation and activity outside of the learning environment, students are still engaged with the environment, if not the content. This may engender more positive feelings towards the environment and also provides opportunities for intervention.

Many hypothesize that the presence of these features is responsible for many of the benefits that game-based learning offers over traditional systems. While removing these features may lead to more on-task learning activities, it may also decrease positive outcomes such as engagement.



Fig. 1. CRYSTAL ISLAND learning environment.

Consequently, it is important to more fully understand how game-features correspond to engagement with the learning tasks.

### 3 CRYSTAL ISLAND

For the past several years, the authors and their colleagues have been designing, implementing, and conducting empirical studies with CRYSTAL ISLAND [36], [45], [46], [47]. CRYSTAL ISLAND (Fig. 1) 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 mystery is derived from the North Carolina state standard course of study for eighth-grade microbiology. CRYSTAL ISLAND’s premise is that a mysterious illness 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. As the student investigates the mystery, she completes an in-game diagnosis worksheet in order to record findings, hypotheses, and a final diagnosis. 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 mystery is solved when the student submits a complete, correct diagnosis and treatment plan to the camp nurse.

To illustrate the behavior of CRYSTAL ISLAND, consider the following situation. Suppose a student has been interacting with the virtual characters in the story world and learning about infectious diseases. In the course of having members of the research team become ill, she has learned that a pathogen is an agent that causes disease in its host and can be transmitted from one organism to another. As the student concludes her introduction to infectious diseases, she uncovers a clue while speaking with a sick patient that suggests the illness may be coming from food items the sick scientists recently ate. Some of the island’s characters

are able to help identify food items and symptoms that are relevant to the scenario, while others are able to provide helpful microbiology information. The student discovers through a series of tests that a container of unpasteurized milk in the dining hall is contaminated with bacteria. By combining this information with her knowledge about the characters’ symptoms, the student deduces that the team is suffering from an *E. coli* outbreak. The student reports her findings back to the camp nurse, and they discuss a plan for treatment.

## 4 CORPUS COLLECTION

As part of an investigation of affect and game-based learning, a study was conducted with 450 eighth grade students from two North Carolina middle schools. Students interacted with the CRYSTAL ISLAND narrative-centered learning environment. After removing instances of incomplete data, the final corpus included data from 400 students. Of these, there were 194 male and 206 female participants. The average age of the students was 13.5 years (SD = 0.62). At the time of the study, the students had not yet completed the microbiology curriculum in their classes.

### 4.1 Method

A week prior to the interaction, students completed a series of pre-study questionnaires including a test of prior knowledge, as well as several measures of personal attributes. Personality was measured using the Big Five Personality Questionnaire, which represents personality along five dimensions: openness, conscientiousness, extraversion, agreeableness and neuroticism [48]. Goal orientation, which refers to the extent that a student values mastery of material and successful performance outcomes when engaged in learning activities, was also measured [49]. Students’ emotion regulation strategies were measured with the Cognitive Emotion Regulation Questionnaire [50], which measures the extent to which each of nine common strategies are used by an individual. Students also completed a researcher-generated curriculum test to assess their domain content knowledge prior to interacting with CRYSTAL ISLAND.

During the study, students interacted with CRYSTAL ISLAND for 55 minutes or until they completed the mystery. During their interaction they received an in-game prompt asking them to report on their cognitive/emotional state at regular seven-minute intervals (Fig. 2). This prompt was described to students as an “experimental social network” that was being pilot tested on CRYSTAL ISLAND. Students selected from one of seven cognitive/emotional states: anxious, bored, confused, curious, excited, focused, and frustrated. This set of states covers both the cognitive and affective phenomena commonly present during learning tasks [22], [23], [24]. They were also asked to type a short status update.” Student actions (dialogue with characters, interactions with objects, etc.) were not interrupted by the notification and they were given up to 1 minute to respond at a convenient time before the notification forced a response. There was no actual cross-student communication enabled by this interface.



Fig. 2. Affect self-report device.

Immediately after completing their interaction with CRYSTAL ISLAND, students were given a post-interaction curriculum test with questions identical to the pre-test. They also completed several questionnaires related to their feelings of immersion and understanding of the CRYSTAL ISLAND mystery. Finally, they completed the Intrinsic Motivation Inventory, which measured their values associated with completing the mystery [51].

## 5 AFFECT AND CRYSTAL ISLAND

Initial work at characterizing the affective experiences of students in CRYSTAL ISLAND (Table 1) has shown some interesting similarities and differences from the results found in other learning environments. For instance, positive, learning-focused affective states such as *focused* (24 percent) and *curious* (19 percent) accounted for the majority of student's self-reported emotions. *Confusion* (16 percent) and *frustration* (16 percent) were the next most frequent emotional states. These states are expected to result from the open-ended nature of the CRYSTAL ISLAND environment. The environment does not tell students specifically what they should be doing at any given time, which may be different from their classroom learning experiences. The somewhat high levels of these emotional states suggest that there may be some students who may need increased levels of guidance, though other students may benefit from exploring the environment on their own. *Excitement* (13 percent) occurred somewhat frequently while highly negative emotions such as *boredom* (8 percent) and *anxiety* (4 percent) were relatively infrequent.

A repeated measure ANOVA indicated that the states occurred at significantly different frequencies,  $F(6, 2394) = 60.4, p < 0.0001$ . Tukey post-hoc analyses indicated the following differences in the occurrence of self-reports:  $focused > curious > (confused = frustrated = excited) > bored > anxious$ . While it is difficult to compare this pattern of states to other environments because of differences in collection procedures, there are interesting similarities and differences. Students interacting with both traditional and game-based environments tend to spend most of their time in a state of *focus* or *engaged concentration* [9]. However, positive affective states such as *delight*, *curiosity* and *excitement*

TABLE 1  
Proportion of Emotion Self-Reports

Emotion	School 1		School 2		Total	
	M	SD	M	SD	M	SD
anxious	4.5%	.083	3.9%	.071	4.3%	.079
bored	8.5%	.165	8.2%	.163	8.4%	.164
confused	16.0%	.166	15.8%	.155	15.9%	.163
curious	18.8%	.147	19.8%	.150	19.1%	.149
excited	13.4%	.163	12.5%	.184	13.1%	.171
focused	22.6%	.192	25.3%	.216	3.5%	.201
frustrated	16.2%	.151	14.5%	.136	15.5%	.143

appear more prevalent in game systems, compared with *confusion* in traditional tutoring systems [9]. These trends suggest that there may be affective benefits to game-based learning environments, though more controlled studies (such as [40]) are necessary to further understand the differences.

### 5.1 Affect and Outcome Measures

Overall, the distribution of affective states found during interactions with the CRYSTAL ISLAND environments suggests that a carefully constructed game-based learning environment may serve to encourage positive affect during learning activities. However, this goal is also motivated by the hypothesis that positive affect is beneficial for both motivational and learning outcomes. Consequently, the next step of analysis investigated how these relationships occur within CRYSTAL ISLAND. First, analyses were conducted to ensure that students experienced learning gains from interacting with CRYSTAL ISLAND. Paired t-tests comparing student's pretest ( $M = 6.6, SD = 2.3$ ) and posttest ( $M = 8.6, SD = 3.4$ ) scores indicated that students' learning gains from using CRYSTAL ISLAND were statistically significant,  $t(399) = 12.5, p < 0.0001$ . Next, correlations were conducted to determine the relationship between the occurrences of these states and learning outcomes. For each affective state the proportion of the student's reports of this state out of the total number of self-reports was correlated with the outcome metrics. Results indicated that positive affect was strongly correlated with learning gains,  $r(398) = 0.16, p = 0.001$ , while negative affect was negatively correlated with learning gains. Additionally, two negatively valenced emotions appeared to be particularly associated with reduced learning. Both *confusion*,  $r(398) = -0.11, p = 0.027$  and *boredom*,  $r(398) = -0.15, p = 0.035$  were negatively correlated with learning outcomes.

Further investigation sought to identify whether students' affective states corresponded with feelings of value, interest, and motivation towards the task. These measures are expected to be positive indicators of student engagement with the activity. Correlations were run between the occurrence of student emotion and five subscales of the Intrinsic Motivation Inventory [51]: Interest/Enjoyment, Perceived Competence, Effort/Importance, Pressure/Tension, and Value/Usefulness. Many affective states had significant correlations with each of these metrics (Table 2). Additionally, there is strong evidence ( $p < 0.001$ ) that positive affect was associated with increased feelings

**TABLE 2**  
Affective States Correlated with Motivation Outcome Measures

	Interest / Enjoyment	Perceived Competence	Effort / Importance	Pressure / Tension	Value / Usefulness
Anxious	0.07	0.00	0.05	0.01	0.09
Bored	<b>-0.43</b>	<b>-0.30</b>	<b>-0.37</b>	0.04	<b>-0.38</b>
Confused	<b>-0.15</b>	-0.11	-0.05	0.10	-0.03
Curious	<b>0.16</b>	0.11	<b>0.13</b>	-0.03	0.08
Excited	<b>0.32</b>	<b>0.22</b>	<b>0.19</b>	-0.08	<b>0.22</b>
Focused	<b>0.22</b>	<b>0.15</b>	<b>0.15</b>	<b>-0.16</b>	<b>0.19</b>
Frustrated	<b>-0.18</b>	-0.09	-0.10	<b>0.15</b>	<b>-0.17</b>

*Bold indicates  $p < 0.05$ , highlighted indicates  $p < 0.01$ .*

of interest ( $r = 0.49$ ), competence ( $r = 0.35$ ), importance ( $r = 0.33$ ) and value ( $r = 0.35$ ). Understandably, positive affect was also associated with reduced feelings of tension ( $r = -0.20$ ).

These results corroborate with many findings in the psychological community which suggest that positive affect is associated with increased learning [2], [3], [5], while negative emotions are believed to lead to decreased motivation and effort [3], [6], [7]. Additionally, there has been strong evidence in the intelligent tutoring system community demonstrating the harmful impacts of negative states, particularly boredom, on learning [9]. However, other researchers have found evidence that confusion may be beneficial for learning [52]. We suspect that these findings were not duplicated in the game-based learning environment because of the highly open-ended nature of the learning experience. It is hypothesized that confusion aids learning when students treat it as a learning opportunity and gain the knowledge that helps them overcome the confusion. [52]. However, we suspect that if the confusion is related specifically to the task, even overcoming this confusion does not aid content learning outcomes. Instead by experiencing confusion of what to do next or how to interact with the environment, the total amount of time students have to interact with learning content is diminished. This is one of the critical open issues in game-based learning as it is important to identify the proper amount of guidance to foster independence but reduce confusion on how to successfully navigate and interact with the environment [41], [53].

### 5.2 Affect and Inquiry Behaviors

The finding that affect was correlated with a variety of outcome measures, including learning gains and feelings of interest, raised the question of how affect may be tied to behaviors in the environment. Prior work examining students' inquiry behaviors in the game-based learning environment suggested that effective inquiry strategies (e.g., gathering background information prior to formulating and testing hypotheses in a virtual laboratory) were not necessarily associated with improved content learning gains [53]. However, effective inquiry strategies were associated with improved problem-solving outcomes.

Conversely, students who did not use good strategies (e.g., gathering background information after formulating and testing hypotheses) were less effective at solving the overall task. These observations led to the hypothesis that individual differences in inquiry strategies may also be associated with differences in affective outcomes. In particular, emotions such as frustration and curiosity were anticipated to correlate with different inquiry behaviors.

For this analysis students were split into two groups based on whether they gathered more information prior to hypothesis testing or waited until they received failed results to gather background information. Overall it was found that more effective inquiry behaviors corresponded with better affective experiences [46]. Specifically, students who performed more information-gathering behaviors prior to hypothesis testing reported more positive emotions, such as curiosity,  $t(318) = 1.97, p = 0.05$ , and excitement,  $t(318) = 2.51, p = 0.01$ . It is possible that these states may have fueled students to learn more about the environment. Good inquiry strategies were also associated with fewer negative cognitive-affective states, like frustration,  $t(122) = 2.09, p = 0.04$ , and confusion,  $t(122) = 2.14, p = 0.03$ . This is likely due to the fact that these students were more efficient at solving the problems. These findings provide further evidence that there is utility in introducing supplemental guidance for some students in game-based learning environments. While some students are able to engage in effective inquiry behaviors and have positive learning and affective outcomes, others are not as successful. Going forward, it will be important to identify and support learners with less effective inquiry strategies to improve the overall experience for all learners.

### 5.3 Affect and Problem Solving

In addition to engaging in inquiry behaviors, students interacting with CRYSTAL ISLAND are engaged in a complex problem-solving task. Findings that good inquiry strategies were correlated with positive affective outcomes prompted an investigation of the relationship between affect, motivation and problem solving. Prior investigations on problem-solving in CRYSTAL ISLAND have investigated various measures of problem solving efficiency such as number of tests run, time to arrive at correct hypotheses, time to solve the mystery, and other similar metrics. However, none of these measures were individually tied to learning gains. Consequently, clustering was used in an attempt to identify patterns of problem solving that are both meaningful and tied to targeted outcomes.

K-means clustering was performed using relevant problem-solving metrics to divide students into 3 clusters. The resulting clusters indicated interesting patterns of problem solving. Cluster 1 ( $N = 110$ ) represented efficient problem solving. These students conducted fewer tests and worksheet checks before arriving at a correct solution. They also moved on to the next problem-solving step after receiving positive feedback. Students in Cluster 2 ( $N = 112$ ) showed puzzling patterns of behavior. They reached positive solutions faster than students in Cluster 1, but would continue testing and gathering information even after they had the solution. This suggests these students may have been

TABLE 3  
Problem Solving Differences

Measure	Cluster 1	Cluster 2	Cluster 3	F	p
Learning gains	.23 (.26)	.19 (.31)	.08 (.13)	10.51	<0.001
Positive affect	.64 (.22)	.56 (.27)	.50 (.25)	9.48	<0.001
Curiosity	.23 (.17)	.17 (.14)	.17 (.16)	4.89	0.008
Competence	4.6 (1.3)	4.9 (1.4)	3.9 (1.5)	19.04	< 0.001
Interest	5.0 (1.5)	4.7 (1.4)	4.3 (1.6)	8.19	< 0.001
Value	5.3 (1.5)	5.0 (1.5)	4.6 (1.6)	5.42	0.005
Effort	5.3 (1.3)	4.9 (1.2)	4.8 (1.3)	4.98	0.007
Boredom	.05 (.11)	.08 (.15)	.11 (.19)	3.95	0.020
Off-task	.03 (.03)	.04 (.06)	.05 (.07)	3.51	0.031

unclear of the next problem solving step, or may have had several possible solutions in mind they wished to investigate. Finally, Cluster 3 ( $N = 178$ ) represents inefficient problem solving. These students took far longer to reach a successful solution and also continued investigations after a positive test, similar to Cluster 2. This pattern of behavior may stem from poor understanding of the problem or a “guess and check” approach.

While the individual metrics that were used to generate the problem-solving clusters did not correlate to learning or engagement outcomes, the resulting clusters had interesting differences between them (Table 3). An ANOVA with Tukey post-hoc tests indicated that students in Cluster 1 and 2 experienced significantly higher learning gains than students in Cluster 3. Furthermore, there were significant differences in measures of affect, engagement, and motivation. Students in Cluster 1 reported positive affective states significantly more than students in Clusters 2 or 3. They also reported feeling more competent, and more interested. Students in Cluster 3 reported more boredom than students in Cluster 1. They also engaged in more off-task and attributed less value to the task. Alternatively, students in Cluster 1 reported more curiosity than students in Clusters 2 or 3 and claimed to put more effort into completing the task.

These findings highlight the importance of a student’s ability to understand the problem solving task and how this ability relates to student affect and engagement. Successful problem solving is tied to increased effort and value, reduced disengagement, and positive affective feelings. Similar to findings related to inquiry strategies, these findings point to the importance of scaffolding problem solving for students who are less successful independently.

#### 5.4 Affect and Off-Task Behaviors

In addition to examining how affect related to positive inquiry and problem solving behaviors, an examination was conducted exploring how affect was tied to students disengaging from the learning task and going off-task. As noted above, we hypothesize that off-task behaviors in a game-based learning environment are inherently different from other types of off-task behaviors in tutorial systems. This is because students may disengage from the learning task, but not from the environment as a whole. By focusing on the game-based aspects of the environment students may be able to “take a break” from difficult learning

activities and receive a boost in positive affect from treating the environment more like a game. It is hoped that this can relieve negative affect and increase the student’s willingness to return to the learning task when ready.

An initial examination of off-task behavior was conducted using students from only the first school ( $N = 260$ ). In this work time spent in non-essential locations or interacting with objects in an unproductive way is classified as off-task [36]. On average, students spent approximately 4.58 percent ( $SD = 6.82$ ) of their time off-task. While this is significantly lower than reported by many other environments (e.g., 18 percent in [26]), it is not unexpected given the very different nature of the environment and unique definition of off-task behavior. There was also a wide range between students with approximately a third of students engaging in no off-task behavior, and at the maximum, one student spent 63.2 percent of his time off-task.

Further analyses were conducted comparing off-task behavior to student learning. Results resembled findings reported from other investigations of off-task behavior in alternate intelligent tutoring systems [26], [30], [31]. Off-task behavior was found to negatively correlate with students’ normalized learning gains,  $r(258) = -0.18, p = 0.004$ . There was no evidence that low prior-knowledge students engaged in more off-task behavior, as the correlation between time off-task and pre-test score was not statistically significant,  $r(258) = 0.08, p = 0.21$ . This result contrasted with a previous investigation of off-task behavior using an earlier version of the CRYSTAL ISLAND learning environment [45].

The results also highlighted evidence that off-task behavior may have a significant affective component. In particular, total time off-task was negatively correlated with *curiosity*  $r(258) = -0.12, p = 0.04$  and *frustration*,  $r(258) = -0.13, p = 0.04$ . This result was surprising given prior work that demonstrated frustration as a trigger for off-task behavior [9]. The finding prompted an examination of whether off-task behavior helps alleviate frustration in the CRYSTAL ISLAND environment.

In order to investigate relationships between off-task behaviors and affect transitions we utilized a measure of transition likelihood,  $L$  (Equation (1)), which calculates the likelihood of a transition between two states relative to chance [24]. The  $L$  statistic has a maximum value of 1, and its minimum value is  $-\infty$ . An  $L$ -value above zero indicates that a particular transition is more likely to occur than chance. A negative  $L$ -value indicates that a state transition is less likely than chance. This statistic is based on Cohen’s kappa, and it is frequently used to measure changes in student emotions that occur over time [9], [24].

To examine whether off-task behavior alleviates frustration or other negative learning emotions, we defined student states as follows: the current state is comprised of the student’s emotion self-report at time  $t_n$  and whether or not the student went off-task between time  $t_n$  and  $t_{n+1}$ . The next state is comprised of the student’s emotion self-report at time  $t_{n+1}$ . The likelihood of transitioning between these two states was calculated using the  $L$  statistic described above. T-tests were conducted to identify transition likelihoods that were significantly different than chance ( $L = 0$ ).

**Equation (1).** *L*-metric

$$L(\textit{Current} \rightarrow \textit{Next}) = \frac{\Pr(\textit{Next}|\textit{Current}) - \Pr(\textit{Next})}{1 - \Pr(\textit{Next})}$$

The analysis revealed that students who reported *frustration* at time  $t_n$  and subsequently went off-task were most likely,  $t(83) = 2.863, p = 0.005$ , to report feeling *focused* seven minutes later at  $t_{n+1}$ . These observations are consistent with a hypothesis that off-task behavior helps to alleviate *frustration*. The finding lends support to the premise that some students use off-task behavior as a way to productively cope with negative affect. A possible explanation for the finding is that students employ emotion self-regulation strategies by taking breaks from challenging tasks, exploring the virtual environment, and returning “refreshed” at later times to re-engage in problem-solving activities. Students who did not go off-task after reporting *frustration* did not appear to reap this same benefit. Frustrated students who stayed on-task were most likely,  $t(140) = 3.43, p < 0.001$ , to report *boredom* at the next self-report. An emotion transition from *frustration* to *boredom* may indicate that a student has disengaged from problem solving altogether.

The hypothesis that students use off-task behavior as a productive strategy for regulating negative affect was not supported when examining affect transitions from the state of *confusion*. Confused students who remained on-task were most likely,  $t(149) = 4.57, p < 0.001$ , to report feeling *focused* at the next self-report, but confused students who went off-task were most likely,  $t(149) = 1.75, p = 0.080$ , to report *boredom*. An affect transition from *confusion* to *boredom* may signify a student reaching an impasse and giving up. These observations also suggest that students who persevered through *confusion* achieved positive affective benefits for doing so. A notable distinction between the *frustration* transitions and the *confusion* transitions is that *frustration* is generally considered harmful for learning, but *confusion* is considered productive for learning despite its negative valence [9].

These findings indicate that off-task behavior is not universally effective for self-regulating negative affect, but the findings also imply that some students may experience emotional benefits from off-task behavior under particular circumstances (e.g., when experiencing frustration). These findings support the hypothesis that off-task behavior may serve a unique purpose in game-based learning environments by offering emotional respite from difficult learning activities but continuing to keep students engaged. To further investigate this correlations were conducted to examine whether students’ off-task behavior was associated with outcome measures of interest and motivation. It was found that students who spent more time off-task were more likely to report lower levels of effort in interacting in the environment  $r(258) = -0.15, p = 0.02$ , which would be expected. However, there was no significant correlation in off-task behavior and measures of interest and enjoyment  $r(258) = -0.04, p = 0.49$  suggesting that the environment still held their attention. This supports the promise of game-based learning environments as a way to support student learning

while encouraging positive affective experiences and engagement with the learning environment. In a follow up investigation that compared students with above-median learning gains and below-median learning gains, the high-learning students tended to transition from frustration to focus after going off-task. On the other hand, students in the low-learning group were no more likely than chance to transition from frustration to focus whether they went off-task or remained on-task. While these findings were not statistically significant, they raised further questions as to whether certain students may experience both learning and affective benefits from going off-task.

## 6 MODELING AFFECT IN CRYSTAL ISLAND

This work on examining the role of affect in game-based learning yielded interesting findings tying affect to learning and motivational outcomes as well as different in-game behaviors. However, these findings also point to the complex nature of the relationship between all these variables. In order to further understand how all the underlying components interact, empirical, machine-learned models of emotion in CRYSTAL ISLAND were investigated.

### 6.1 Empirically Learned Models

Because of the inherent uncertainty in modeling student emotion, Bayesian networks were used for predicting student affective states. Bayesian networks are graphical models used to model processes under uncertainty by representing the relationship between variables in terms of a probability distribution [54]. Another reason for selecting a Bayesian representation was the ability to incorporate both theoretical and empirical knowledge. Bayesian networks involve two main components, (1) a network structure, which describes which variables are related to others, and (2) a set of conditional dependencies which provide the exact specifications for these relationships. Both the structure and the conditional dependencies can be learned using a variety of possible algorithms [55] or specified by hand.

The first area of work that was pursued was determining whether theoretical models of learner emotions could be used to guide the development of empirical models [47]. Therefore the Bayesian networks were designed with the structure informed by the proposed relationships described within Elliot and Pekrun’s [56] model of learner emotions as well as prior findings on the role of off-task behavior for regulating affect. Each Bayesian network was specified using the GeNIe modeling environment developed by the Decision Systems Laboratory of the University of Pittsburgh (<http://dsl.sis.pitt.edu>). After hand-crafting the structure of the Bayesian network, the parameters were learned using the EM algorithm provided by GeNIe.

The models contained three types of variables:

1. *Personal attributes*. These static attributes were taken directly from students’ scores on the personal surveys prior to the interaction. Included were all four attributes for goal orientation and three personality attributes expected to be relevant to the

student's appraisal: conscientiousness, openness, and agreeableness.

2. *Observable environment variables.* These dynamic attributes capture a snapshot of the student's activity in the learning environment up until the time of the self-report. They provide a summary of important actions taken, such as TestsRun, BooksViewed, and GoalsCompleted. They also include information about how well the student is doing in the environment based on certain milestones, such as SuccessfulTest and WorksheetChecks. Two features measuring the student's level of off-task behavior were also included because of their hypothesized role in emotion regulation.
3. *Appraisal variables.* The values of the appraisal variables are not directly observable in the environment. Instead they are the result of the student's cognitive appraisal of many factors. The selected appraisal variables and their relation to observable variables are informed by the model of learner emotions.

Each learned model was evaluated using student-level 10-fold cross-validation. In this technique, a model's parameters are trained using data from 90 percent of the students. The predictive accuracy of the model is then evaluated on the remaining 10 percent of the corpus. This approach is designed to provide an accurate measure of how well a trained model will extend to future, unseen populations.

The learned models were evaluated on their ability to correctly predict both the specific emotion label as well as the valence (positive or negative) of the emotional state. A baseline measure of most-frequent class offered an accuracy of 22.4 percent for emotion prediction and 54.5 percent for valence prediction. These levels offer a more conservative estimate than a random model, which would have baselines of 14.3 and 50 percent respectively.

In order to provide an additional baseline of comparison, a naïve Bayesian network was learned. A naïve Bayesian network operates under the "naïve" assumption that all variables are directly related to the outcome variable but are conditionally independent of each other [55]. The learned naïve Bayesian network achieved a predictive accuracy of 18.1 percent on emotion label and 51.2 percent on valence. This performance is less accurate than the most frequent label baseline model, but provides an additional baseline measure. By comparing carefully constructed Bayesian networks against the naïve assumption we can determine the degree to which affective models benefit from theoretically informed structure.

## 6.2 Bayesian Network

First, a Bayesian network (Fig. 3) was designed with the structure informed by the proposed relationships described within Elliot and Pekrun's model of learner emotions [47], [56]. The design of the structure focused on the appraisal of learning and performance goals and how these goals were being met based on the status of the game environment. For example, learning-focused activities such as book reading or note-taking are expected to impact how much a student's learning goals are being met, while performance appraisals are more likely related to achieving important milestones

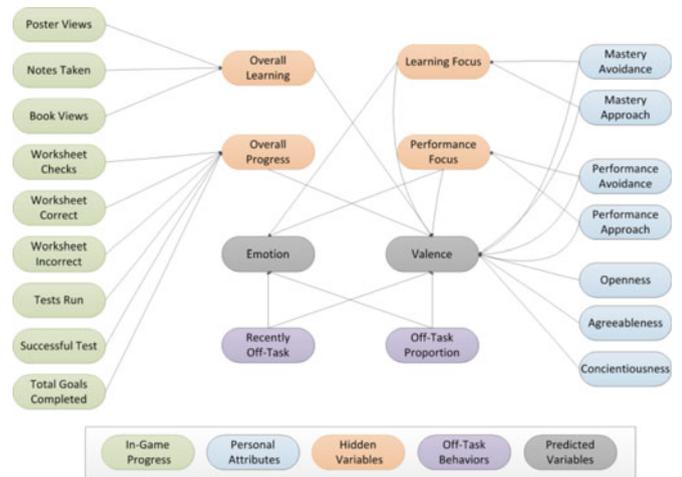


Fig. 3. Structure of Bayesian network.

such as running a successful test. Meanwhile, goal focus and valence tendencies are considered to be dependent on their personal attributes as described by the model. For example, students with approach orientations are expected to have generally more positive temperaments and emotional experiences than students with avoidance orientations. Similarly, personality traits such as agreeableness and openness are expected to contribute to overall temperament.

The evaluation of the static Bayesian network showed that it was able to predict the emotion label with 26.3 percent accuracy and the valence with 70.1 percent accuracy. Both of these predictions offer a significant gain over the most frequent baseline and the naïve Bayesian network ( $p < 0.05$ ). This improvement highlights the benefits of using a theoretical model of learner emotions to guide the model's structure.

However, this model does not capture the dynamic nature of emotions over time and specifically does not take into account the previous emotion self-reports. To account for this the static Bayesian network was extended into a dynamic Bayesian network. Dynamic Bayesian Networks are able to account for temporal relationships between variables, allowing observations at time  $t_n$  to inform observations at time  $t_{n+1}$ . Utilizing this framework, we extended the static Bayesian network to include temporal relationships between valence and emotion across time. Again, the model was trained using GeNIe's EM algorithm and evaluated with 10-fold cross-validation. This model was able to offer further improvements over the static Bayesian network. It predicted emotion with 34.7 percent accuracy and valence with 80.6 percent accuracy.

While these predictive accuracies are significantly greater than baseline the accuracy for predicting emotion labels is still particularly low. This is likely due to the large number of affect labels and possibly the similarity of the states being predicted. For example, the states of *confusion* and *frustration* have very similar qualities and often precede one another [9], so it is unsurprising that a machine-learned model has difficulties distinguishing the two. Further work is necessary to develop models better able to distinguish emotion labels and identify the predictive accuracy that is necessary to inform an affect-sensitive system.

### 6.3 Affective Reasoning

Beyond predicting affective states, these learned models of affect have also been used to reason about the outcomes of student behaviors. Prompted by earlier findings that off-task behavior may be used to regulate affect, investigations were conducted to see if these models could be utilized to identify cases where off-task behavior had a positive affective outcome [57]. To this end the DBN was used to generate alternate futures that simulated students' affective trajectories as if they had performed fewer off-task behaviors than in reality. The alternate futures were compared to students' actual affective trajectories in order to generate labels indicating whether off-task behaviors were cases of emotion self-regulation.

These labels were then used to divide students into a group who appeared to be capable of using off-task behavior as a productive means of reducing negative affect and those who may not have this skill. Of the 260 students in the collected corpus, 68 students did not engage in any off-task behavior so their data was removed from this portion of the analysis. By comparing the real student logs to the proposed alternate outcomes, it was suggested that 78 students were reaping a positive affective outcome from off-task behavior and the remaining 114 did not receive the same benefit. Prior evidence suggested that students who can better regulate their affect have more successful learning outcomes.

Analyses comparing these two groups found that the average learning gains of the group of students who were not using off-task behavior to regulate affect was 1.21 ( $SD = 3.24$ ). Meanwhile the learning gains for the students who evidence affect regulation was 2.40 ( $SD = 3.29$ ). T-tests indicate that this is a significant difference in learning gains,  $t(190) = 2.50, p = 0.013$ , with regulating students experiencing nearly double the learning gains of the non-regulating group. This finding is particularly interesting since it suggests that the students who used off-task behavior to regulate their affect were less likely to have experienced harmful impacts on their learning gains. In fact, while there is a strong correlation between off-task behavior and learning gains,  $r(112) = -0.23, p = 0.015$ , for the students who did not demonstrate evidence of emotion regulation, there is no significant correlation for the students in the affect regulation group,  $r(76) = -0.06, p = 0.587$ . These findings indicate that this methodology has promise in identifying students who are able to effectively regulate their affective states.

## 7 DISCUSSION

The importance of affect and engagement in supporting student learning has been demonstrated in a multitude of computer-based learning environments. This work examined these relationships within the game-based learning environment, CRYSTAL ISLAND. Game-based learning is the subject of increasing attention, in large part because of hypothesized increases in motivation, interest, and ultimately, learning, so it is important to examine how these aspects of student experience actually occur.

Overall, the examination of affect, engagement and learning in CRYSTAL ISLAND showed that students were reporting more positive affect experiences than negative. Students are particularly less likely to report highly negative emotions

such as boredom and anxiety though these states are often reported during other learning activities. As in other learning systems positive affect was correlated with increased learning. It was also tied to increased measures of interest and more on-task behavior. Analysis of key strategies related to inquiry and problem solving showed positive correlations with affect and engagement as well. Students who exhibited these strategies had better affective outcomes and attributed more value and interest to the activity. Overall, positive affect was found to have the same ties to learning behaviors and outcomes as is reported in many traditional tutorial systems.

However, one particularly interesting finding was a difference in the role of confusion as it relates to learning outcomes. In many systems occurrences of confusion are correlated with learning, as the very experience of confusion offers an opportunity to master a previously unknown concept. However, this was not the case in CRYSTAL ISLAND. One hypothesis is that the open-ended nature of a game-based learning environment may introduce additional sources of confusion and that these experiences do not tie to learning gains but instead reduce opportunities for students to learn the material. Results showing the differences in problem-solving abilities support the idea that some students struggle with this task though there were no direct ties to confusion. These findings suggest that striking the right balance between independent and guided problem solving is still a major open issue in game-based learning [41] and is something that will need to be examined further within CRYSTAL ISLAND.

Another primary concern of game-based learning environments is the existence of superfluous features that may distract students from the learning task. However, the occurrence of off-task behavior in CRYSTAL ISLAND was not found to be exceedingly high and is comparable to reports in other learning systems. Similarly, we also found off-task behavior to be negatively correlated with learning gains and particularly linked to negative affect. However, it was found that off-task behavior may aid students in regulating negative affect and returning to the learning task in a more productive frame of mind. This finding suggests that these extraneous features of game-based learning environments are not always harmful and should not simply be removed. Instead it points to the need of an adaptive system which can encourage students to return to the task when their non-learning play does not appear to be productive.

Models of student affect and engagement were examined to begin to address these very problems. Results suggest that there is significant promise in the use of these models to predict both how a student is feeling and what the affective consequences of certain behaviors may be. These models could prove a powerful tool in providing adaptive scaffolding to support student affect.

## 8 LIMITATIONS

While this work highlights many interesting relationships between affect, engagement and learning in game-based environments, there are several limitations to consider. First, the work relies on self-reports for identification of affective states. The interpretation of these results relies

upon the assumption that students were honest and accurate with the identification of their affective states. The framing as a social network may also skew the affective states more positively due to social desirability effects. Alternative approaches to identifying affect include using trained human judges or automatic facial feature-based affect detection. However, each of these approaches has tradeoffs as there is still no clear gold-standard for emotion recognition. Beyond ease of collection, there are further benefits of choosing self-report for affect labeling. Specifically, we hope to investigate affective feedback in response to students' self-reported emotions. We expect that responding to the user's reported state will reduce the likelihood that the student is jarred by affective feedback. The student will know how the game received affective information and consequently there is reduced risk of incorrect labeling and responding to a state that the user is not feeling. Furthermore, the social network self-report tool provides an interesting avenue for delivering feedback.

Another limitation related to the self-report approach involves the duration of time that passes between each report. Many different affective states can occur over a duration of 7 minutes making it difficult to interpret time-sensitive results. This duration was chosen so as to not annoy students with constant interruptions. Automated techniques or human judges may allow better granularity of analyses in the future. However, we hope that in the current study students report the general affective state they have been feeling over that time. While this granularity may not capture some moment-by-moment affective qualities, we expect that we have a general understanding of how the student has been feeling over time.

Students were also required to select one of the provided affective states. There was no option for neutral or no response. They also could not identify their own state or "other." While it is possible that we are missing out on some rich affective experiences, we chose states that have been reported with enough frequency to be analyzed in the past. We also chose not to include neutral or no state to simplify data analyses and to encourage students to pick the state that best described their current emotion. Furthermore, the focused state is likely very similar to a neutral state given the problem-solving task. This may describe the relatively high occurrence of this state compared with others. Since the states were not specifically described to students and were based on personal interpretations of the labels it would be worthwhile to examine what valence the students attributed to this state.

## 9 FUTURE DIRECTIONS

These findings present many interesting opportunities for future work. First among these is further investigation into the role of confusion and cognitive load in open-ended game-based learning environments. Since this finding is one that directly contradicted findings in traditional tutorial environments it is important to further examine how confusion and learning occur within *CRYSTAL ISLAND*. This line of work will likely involve identifying different possible sources of confusion in order to understand which are most harmful.

Along this direction it will also be important to develop methods that scaffold learning and affect within the environment. Off-task behavior is a way that some students can regulate their own affect, but not all students are able to do this. Current models can endeavor to detect when students should be allowed to go off-task and which should be encouraged to return to a learning focus. However, it will also be important to examine how additional mechanisms of scaffolding affect, such as empathetic agents or content hints or guidance, impact learning and engagement [58].

Finally, an important future direction is to provide direct comparison between game-based learning environments and traditional tutorial systems. It will be important to develop a methodology of comparing these systems that is not biased due to time constraints, hastily developed game features, or other confounding factors. Identifying the necessary requirements to conduct such an investigation is just as important as examining the findings that arise. Only with such a carefully crafted examination is it possible to understand the similarities and differences between traditional and game-based learning environments.

## 10 CONCLUSION

Affect permeates every aspect of human experience, including learning. The types of emotions we associate with learning influence how we perceive learning and how likely we are to actively engage in learning activities. Because of this significant impact on learning, understanding the role of affect and engagement in computer-based learning environments is the subject of increasing attention. While game-based approaches show great promise for increasing engagement, positive affect, and learning, there still remain outstanding issues of how game features may influence learners' experiences.

This work has explored the role of affect and engagement in the game-based learning environment, *CRYSTAL ISLAND*. Results highlight the relationships between students' emotions, motivations, interest, and learning outcomes. Specifically, positive emotions are associated with increased learning and motivation. Students with better inquiry and problem-solving skills tend to have better affective outcomes when interacting with the open-ended environment. Meanwhile, disengagement and off-task behavior is associated more closely with negative affective states, but may be used as a means of emotion regulation. Empirical models of affect and disengagement suggest that students who use off-task behavior for emotion regulation do not suffer the negative learning outcomes typically associated with disengagement.

This work highlights the strong relationships between cognitive and affective phenomena in game-based learning environments and calls for the systematic exploration of the role of emotion in learning. Future work on affect in game-based learning should investigate how emotion regulation modulates these processes. Meanwhile, controlled comparisons between game-based and traditional learning environments will help to empirically identify the role of games in supporting affective experiences of learners. With these results in hand, the field is poised to explore the many open questions on how games and game-like

environments can best be utilized to support affect and engagement during learning.

## 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 US National Science Foundation (NSF) 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 NSF. Additional support was provided by the Bill and Melinda Gates Foundation, the William and Flora Hewlett Foundation, and EDUCAUSE.

## REFERENCES

- [1] J. Forgas, C. Wyland, and S. Lahan, "Hearts and Minds: An Introduction to the Role of Affect in Social Cognition and Behavior," *Affect in Social Thinking and Behavior*, J.P. Forgas, ed., pp. 3-18, Psychology Press, 2006.
- [2] R. Kanfer and P.L. Ackerman, "Motivation and Cognitive Abilities: An Integrative/Aptitude-Treatment Interaction Approach to Skill Acquisition," *J. Applied Psychology*, vol. 74, pp. 657-690, 1989.
- [3] R. Pekrun, T. Goetz, W. Titz, and R. Perry, "Academic Emotions in Students' Self-Regulated Learning and Achievement: A Program of Qualitative and Quantitative Research," *Educational Psychologist*, vol. 37, no. 2, pp. 91-105, June 2002.
- [4] R. Raghunathan and Y. Trope, "Walking the Tightrope between Feeling Good and Being Accurate: Mood as a Resource in Processing Persuasive Messages," *J. Personality and Social Psychology*, vol. 83, pp. 510-525, 2002.
- [5] H. Bless, G. Clore, N. Schwarz, V. Golisano, C. Rabe, and M. Wolk, "Mood and the Use of Scripts: Does a Happy Mood Really Lead to Mindlessness?" *J. Personality and Social Psychology*, vol. 71, pp. 665-679, 1996.
- [6] D.K. Meyer and J.C. Turner, "Re-Conceptualizing Emotion and Motivation to Learn in Classroom Contexts," *Educational Psychology Rev.*, vol. 18, no. 4, pp. 377-390, Oct. 2006.
- [7] O.M. Ramirez and C.J. Dockweiler, "Mathematics Anxiety: A Systematic Review," *Advances in Test Anxiety Research*, R. Schwarzer, H.M. Ploeg, and C.D. Spielberger, eds., pp. 157-175, Erlbaum Assoc., 1987.
- [8] R.W. Picard et al., "Affective Learning—A Manifesto," *BT Technology J.*, vol. 22, no. 4, pp. 253-269, Oct. 2004.
- [9] R. S. Baker, S. D'Mello, S. K. Rodrigo, and A.C. Graesser, "Better to Be Frustrated than Bored: The Incidence, Persistence, and Impact of Learners' Cognitive-Affective States during Interactions with Three Different Computer-Based Learning Environments," *Int'l J. Human-Computer Studies*, vol. 68, no. 4, pp. 223-241, 2010.
- [10] J. Sabourin, J.P. Rowe, B.W. Mott, and J.C. Lester, "When Off-Task Is On-Task: The Affective Role of Off-Task Behavior in Narrative-Centered Learning Environments," *Proc. 15th Int'l Conf. Artificial Intelligence in Education*, pp. 523-536, 2011.
- [11] S. McQuiggan, J. Robison, and J. Lester, "Affective Transitions in Narrative-Centered Learning Environments," *Educational Technology & Soc.*, vol. 13, no. 1, pp. 460-464, 2010.
- [12] K. Millis, C. Forsyth, H. Butler, P. Wallace, A. Graesser, and D. Halpern, "Operation ARIES! a Serious Game for Teaching Scientific Inquiry," *Serious Games and Edutainment Applications*, M. Ma, A. Oikonomou, and L.C. Jain, eds., pp. 169-195, Springer-Verlag, 2011.
- [13] C. Conati, "Probabilistic Assessment of User's Emotions in Educational Games," *J. Applied Artificial Intelligence, Special Issue on Merging Cognition and Affect in HCI*, vol. 16, no. 7-8, pp. 555-575, 2002.
- [14] J.P. Rowe, L.R. Shores, B.W. Mott, and J.C. Lester, "Integrating Learning, Problem Solving, and Engagement in Narrative-Centered Learning Environments," *Int'l J. Artificial Intelligence in Education*, vol. 21, pp. 115-133, 2011.
- [15] D.W. Shaffer, *How Computer Games Help Children Learn*. Palgrave Macmillan, 2006.
- [16] J.P. Gee, *What Video Games Have to Teach Us about Learning and Literacy*. Palgrave Macmillan, 2003.
- [17] D.S. McNamara, G.T. Jackson, and A.C. Graesser, "Intelligent Tutoring and Games (ITaG)," *Proc. Workshop Intelligent Educational Games at the 14th Int'l Conf. Artificial Intelligence in Education*, pp. 1-10, 2009.
- [18] K. Kapp, *The Gamification of Learning and Instruction*. Pfeiffer, 2012.
- [19] B.S. Bloom, "The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring," *Educational Researcher*, vol. 13, no. 6, pp. 4-16, 1984.
- [20] K. VanLehn, "The Behavior of Tutoring Systems," *Int'l J. Artificial Intelligence in Education*, vol. 16, no. 3, pp. 227-265, 2006.
- [21] B.P. Woolf, *Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing e-Learning*. Morgan Kaufmann, 2009.
- [22] B. Kort, R. Reilly, and R. Picard, "An Affective Model of Interplay Between Emotions and Learning: Reengineering Educational Pedagogy—Building a Learning Companion," *Proc. IEEE Int'l Conf. Advanced Learning Technology*, 2001.
- [23] R. Baker, M. Rodrigo, and U. Xolocotzin, "The Dynamics of Affective Transitions in Simulation Problem-Solving Environments," *Proc. Second Int'l Conf. Affective Computing and Intelligent Interactions*, pp. 666-677, 2007.
- [24] S. D'Mello, R.S. Taylor, and A.C. Graesser, "Monitoring Affective Trajectories during Complex Learning," *Proc. 29th Ann. Meeting of the Cognitive Science Soc.*, pp. 203-208, 2007.
- [25] K. Muldner, W. Burselen, B. Van de Sand, and K. VanLehn, "An Analysis of Gaming Behaviors in an Intelligent Tutoring System," *Proc. 10th Int'l Conf. Intelligent Tutoring Systems2*, pp. 184-193, 2010.
- [26] R.S. Baker, A.T. Corbett, K.R. Koedinger, and A. Wagner, "Off-Task Behavior in the Cognitive Tutor Classroom: When Students Game the System," *Proc. SIGCHI Conf. Human Factors in Computing Systems*, pp. 383-390, 2004.
- [27] C.R. Beal, L. Qu, and H. Lee, "Classifying Learner Engagement through Integration of Multiple Data Sources," *Proc. Nat'l Conf. Artificial Intelligence*, pp. 2-8, 2006.
- [28] I. Arroyo et al., "Repairing Disengagement with Non-Invasive Interventions," *Proc. 13th Int'l Conf. Artificial Intelligence in Education*, pp. 195-202, 2007.
- [29] R.S. Baker, "Modeling and Understanding Students' Off-Task Behavior in Intelligent Tutoring Systems," *Proc. SIGCHI Conf. Human Factors in Computing Systems*, pp. 1059-1068, 2007.
- [30] M. Cocea, A. Hershkovitz, and R.S. Baker, "The Impact of Off-Task and Gaming Behaviors on Learning: Immediate or Aggregate," *Proc. 14th Int'l Conf. Artificial Intelligence in Education*, pp. 507-514, 2009.
- [31] Y. Gong, J. Beck, N. Heffernan, and E. Forbes-Summers, "The Fine-Grained Impact of Gaming (?) on Learning," *Proc. 10th Int'l Conf. Intelligent Tutoring Systems*, pp. 194-203, 2010.
- [32] R.S. Baker, A.T. Corbett, I. Roll, and K.R. Koedinger, "Developing a Generalizable Detector of When Students Game the System," *User Modeling and User-Adapted Interaction*, vol. 18, no. 3, pp. 287-314, 2008.
- [33] S. Cetintas, S. Luo, Y. Xin, C. Hord, and D. Zhang, "Learning to Identify Students' Off-Task Behavior in Intelligent Tutoring Systems," *Proc. 14th Int'l Conf. Artificial Intelligence in Education2*, pp. 701-703, 2009.
- [34] C.R. Beal and L. Qu, "Relating Machine Estimates of Students' Learning Goals to Learning Outcomes: A DBN Approach," *Proc. 13th Int'l Conf. Artificial Intelligence in Education*, pp. 111-118, 2007.
- [35] R.S. Baker et al., "The Dynamics between Student Affect and Behavior Occurring Outside of Educational Software," *Proc. Fourth Int'l Conf. Affective Computing and Intelligent Interaction*, pp. 14-24, 2011.
- [36] J.L. Sabourin, J.P. Rowe, B.W. Mott, and J.C. Lester, "When Off-Task is On-Task: The Affective Role of Off-Task Behavior in Narrative-Centered Learning Environments," *Proc. 15th Int'l Conf. Artificial Intelligence and Education*, pp. 534-536, 2011.
- [37] D.J. Ketelhut, "The Impact of Student Self-Efficacy on Scientific Inquiry Skills: An Exploratory Investigation in 'River City', a Multi-User Virtual Environment," *J. Science Education and Technology*, vol. 16, no. 1, pp. 99-111, 2007.

- [38] J. Kim et al., "BiLAT: A Game-Based Environment for Practicing Negotiation in a Cultural Context," *Int'l J. Artificial Intelligence in Education*, vol. 19, no. 3, pp. 289-308, 2009.
- [39] W.L. Johnson, "Serious Use of a Serious Game for Language Learning," *Int'l J. Artificial Intelligence in Education*, vol. 20, no. 2, pp. 175-195, 2010.
- [40] N. Hallinen, E. Walker, R. Wylie, A. Ogan, and C. Jones, "I Was Playing When I Learned: A Narrative Game for French Aspectual Distinctions," *Proc. Workshop Intelligent Educational Games at the 14th Int'l Conf. Artificial Intelligence in Education*, pp. 117-120, 2009.
- [41] M.W. Easterday, V. Aleven, R. Scheines, and S.M. Carver, "Using Tutors to Improve Educational Games," *Proc. 15th Int'l Conf. Artificial Intelligence in Education*, pp. 63-71, 2011.
- [42] P. Wouters, C. van Nimwegen, H. van Oostendorp, and E.D. van der Spek, "A Meta-Analysis of the Cognitive and Motivational Effects of Serious Games," *J. Educational Psychology*, vol. 105, no. 2, pp. 248-265, 2013.
- [43] R.E. Mayer and C.I. Johnson, "Adding Instructional Features that Promote Learning in a Game-Like Environment," *J. Educational Computing Research*, vol. 42, no. 3, pp. 241-265, 2010.
- [44] S. Harp and R.E. Mayer, "How Seductive Details Do Their Damage: A Theory of Cognitive Interest in Science Learning," *J. Educational Psychology*, vol. 90, no. 3, pp. 414-434, 1998.
- [45] J.P. Rowe, S.W. McQuiggan, J.L. Robison, and J.C. Lester, "Off-Task Behavior in Narrative-Centered Learning Environments," *Proc. 14th Int'l Conf. Artificial Intelligence and Education*, pp. 99-106, 2009.
- [46] J. Sabourin, J.P. Rowe, B.W. Mott, and J.C. Lester, "Exploring Affect and Inquiry in Open-Ended Game-Based Learning Environments," *Proc. Workshop Emotions in Games for Learning in Conjunction with the 11th Int'l Conf. Intelligent Tutoring Systems*, 2012.
- [47] J. Sabourin, B. Mott, and J. Lester, "Modeling Learner Affect with Theoretically Grounded Dynamic Bayesian Networks," *Proc. Fourth Int'l Conf. Affective Computing and Intelligent Interaction*, pp. 286-295, 2011.
- [48] R. McCrae and P. Costa, *Personality in Adulthood: A Five-Factor Theory Perspective*, second ed. Guilford Press, 1993.
- [49] A. Elliot and H.A. McGregor, "A 2 x 2 Achievement Goal Framework," *J. Personality and Social Psychology*, vol. 80, pp. 501-519, 2001.
- [50] N. Gernefski and V. Kraati, "Cognitive Emotion Regulation Questionnaire: Development of a Short 18-Item Version," *Personality and Individual Differences*, vol. 41, pp. 1045-1053, 2006.
- [51] E. McAuley, T. Duncan, and V. Tammen, "Psychometric Properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting: A Confirmatory Factor Analysis," *Research Quarterly for Exercise and Sport*, vol. 60, no. 1, pp. 48-58, 1989.
- [52] S.K. D'Mello, B. Lehman, R. Pekrun, and A.C. Graesser, "Confusion Can Be Beneficial for Learning," *Learning and Instruction*, 2012.
- [53] J. Sabourin, J.P. Rowe, B.W. Mott, and J.C. Lester, "Exploring Inquiry-Based Problem-Solving Strategies in Game-Based Learning Environments," *Proc. 11th Int'l Conf. Intelligent Tutoring Systems*, pp. 470-475, 2012.
- [54] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, second ed. Pearson, 2003.
- [55] E. Alpaydin, *Introduction to Machine Learning*. MIT Press, 2004.
- [56] A. Elliot and R. Pekrun, "Emotion in the Hierarchical Model of Approach-Avoidance Achievement Motivation," *Emotion in Education*, P. Schutz and R. Pekrun, eds., pp. 57-74, Elsevier, 2007.
- [57] J.L. Sabourin, J.P. Rowe, B.W. Mott, and J.C. Lester, "Considering Alternate Futures to Classify Off-Task Behavior as Emotion Self-Regulation: A Supervised Learning Approach," *J. Educational Data Mining*, vol. 5, no. 1, pp. 9-38, 2013.
- [58] J.L. Robison, S.W. McQuiggan, and J.C. Lester, "Evaluating the Consequences of Affective Feedback in Intelligent Tutoring Systems," *Proc. Int'l Conf. Affective Computing & Intelligent Interaction*, pp. 37-42, 2009.

**Jennifer L. Sabourin** received the BS and MS degrees in 2008 and 2012, respectively, in computer science from North Carolina State University, where she graduated as Valedictorian. She is currently working toward the PhD degree in the Department of Computer Science at North Carolina State University and was a research scientist and a software developer at SAS Institute. She received the US National Science Foundation Graduate Research Fellowship award.

**James C. Lester** received the BA degree in history in 1983 from Baylor University, and the BA degree in 1986, the MSCS degree in 1988, and the PhD degree in 1994 in computer science from the University of Texas at Austin. A member of Phi Beta Kappa, he has been recognized with the US National Science Foundation (NSF) CAREER Award and several Best Paper Awards. He has served as a program chair for the ACM Conference on Intelligent User Interfaces, the International Conference on Intelligent Tutoring Systems, and the International Conference on Foundations of Digital Games. He has served as a conference co-chair for the International Conference on Intelligent Virtual Agents and an editor-in-chief for the *International Journal Artificial Intelligence in Education*.

► **For more information on this or any other computing topic, please visit our Digital Library at [www.computer.org/publications/dlib](http://www.computer.org/publications/dlib).**