Affect and Usage Choices in Simulation Problem-Solving Environments

Ma. Mercedes T. RODRIGO[1], Ryan S.J.d. BAKER[2], Maria C.V. LAGUD[3], Sheryl A.L. LIM[1], Alexis F. MACAPANPAN[3], Sheila A.M.S. PASCUA[3], Jerry Q. SANTILLANO[1], Leima R.S. SEVILLA[3], Jessica O. SUGAY[1], Sinath TEP[1], Norma J.B. VIEHLAND[3]

[1] Department of Information Systems and Computer Science, Ateneo de Manila University
[2] Learning Sciences Research Institute, University of Nottingham
[3] Education Department, Ateneo de Manila University

mrodrigo@ateneo.edu, ryan@educationaldatamining.org

Abstract. We investigate the relationship between a student's affect and how he or she chooses to use a simulation problem-solving environment, using quantitative field observations. Within the environment studied, many students were observed gaming the system (cf. Baker et al, 2004), while few students engaged in off-task behavior. We analyze which affective states co-occur with gaming the system, and which affective states precede gaming behavior. Boredom and confusion appear both to precede gaming behavior and to co-occur with gaming behavior; delight and flow are negatively associated with gaming behavior.

1. Introduction

In recent years, there has been increasing interest in understanding how a student's attitudes and affective state concretely alters his or her behavior, as he or she uses an interactive learning environment [1,5,6,8,16]. Much of this research was conducted by assessing each student's general affect and/or attitudes towards the learning environment through questionnaires given before or after system usage. Then, data mining is used to link each student's behavior to his or her self-reported attitudes and affect [1,5,6,16]. Studies following this approach have produced understanding of the links between affect, attitudes, and students' actions in learning systems. The majority of this work, however, has only studied the links between a student's generalized affect towards a system and his or her behavior, rather than the relationship between a student's affective states at a specific moment, and his or her behavior at the same time or immediately afterwards. (One exception, [8], explicitly models student affect at specific times, but represents affect solely as an outcome of the interaction between the student and the system, rather than as a factor which potentially alters student behavior).

Studying affect only in general, rather than at specific times, limits the questions that an analysis of the relationship between affect and behavior can address. To give a pair of examples: Even if we have evidence that gaming the system is associated with frustration [cf. 16], or that taking a long time between answering attempts is generally linked with fear of making errors [cf. 1], does that mean that a student first experiences the affective state (frustration, fear), and then engages in the behavior? Or does the student experience the affective state as he or she is engaging in the behavior? Or is the relationship between affect and behavior more complex still? Questionnaire assessments of affect and attitudes are an important beginning towards showing the relationship between affective states and behavior in learning systems, but leave some important questions unanswered.

In this paper, we analyze the relationship between a student's affective state, at a given time, and their behavior, both at that time and shortly thereafter. In this manner, we can study more precisely how specific affective states are antecedents to and/or co-occur with a student's choices of how to interact with a learning system.

More specifically, we will study two categories of behavior found to be associated with poorer learning – gaming the system [cf. 4], and off-task behavior [2]. Prior research has shown that gaming the system is generally associated with frustration [16], but it is not yet clear whether students experience frustration while gaming or whether students experience frustration and then game the system shortly afterwards. Additionally, both gaming the system and off-task behavior have been found to be associated with lack of interest in the system's subject matter [2, 16] – hence, boredom may also be associated with gaming the system and off-task behavior.

In this paper, we study the relationship between affect and student usage choices within a simulation problem solving environment designed to be both fun and educational, The Incredible Machine: Even More Contraptions [15]. Studying this issue within the context of a simulation problem solving environment enables us both to learn about the relationships between specific behaviors and affective states, and to learn how the frequencies of behaviors and affective states differ between simulation learning environments and the other types of environments where these behaviors and affective states have been studied [cf. 4, 9, 17]. In particular, it has been found that off-task behavior is considerably less common in action games than in intelligent tutoring systems [cf. 4, 17]. By seeing whether the incidence of off-task behavior in a simulation learning environment designed to be both fun and educational is more similar to the incidence of off-task behavior in action games – systems designed primarily with the goal of fun – or intelligent tutoring systems – systems designed primarily with the goal of learning, we can understand better how students view learning environments designed with both goals in mind.

2. Methods

The relationship between affective states and usage choices was studied within a high school mathematics class in a private school in urban Manila, in the Philippines. Student ages ranged from 14 to 19, with an average age of 16. Thirty-six students participated in this study (17 female, 19 male).

Each student used The Incredible Machine: Even More Contraptions [15] (shown in Figure 1), a simulation environment where the user completes a series of logical "Rube Goldberg" puzzles. In each puzzle, the student has a pre-selected set of

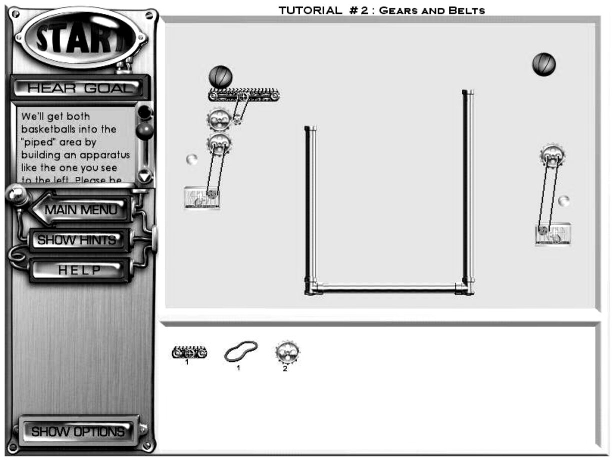

Figure 1. A screen shot from The Incredible Machine: Even More Contraptions.

objects to use, such as scissors, ropes, and pulleys, electrical generators, and animals. The student must combine these objects in order to accomplish a pre-defined goal, such as lighting a candle or making a mouse run. If a student is stuck, he or she can ask for a hint; hint messages display where items should be located in a correct solution to the current problem (but do not show which item should be placed in each location).

Each student used The Incredible Machine for ten minutes, and each student's behavior and affect was observed several times as he or she used The Incredible Machine. The observations were conducted using a method which incorporated aspects of Baker et al's [4] quantitative field observations of student behavior categories, and Craig et al's [9] laboratory observations of affect. The observations were carried out by a team of six observers, working in pairs. As in Baker et al, each observation lasted twenty seconds, and was conducted using peripheral vision in order to make it less clear exactly when an observation was occurring. If two distinct behaviors were seen during an observation, only the first behavior observed was coded, and any behavior by a student other than the student currently being observed was not coded.

It was not possible for the entire class to use the software at the same time, due to the size of the school computer laboratory; hence, students used the software in groups of nine (one student per computer), during their class time. Each pair of observers was assigned to three students and alternated between them. Since each observation lasted twenty seconds, each student was observed once per minute. Observing students more frequently than in [4] or [9] made it possible to directly analyze the relationship between a student's affective state at a given time and their usage choices shortly thereafter.

Within an observation, each observer coded both the student's behavior and affective state, using coding schemes developed in prior research. The observers trained for the task through a series of pre-observation discussions on the meaning of the usage and affective categories.

The usage categories coded were adapted from [4], and are as follows:

1. **On-task** – working within The Incredible Machine
2. **On-task conversation** – talking to the teacher or another student about The Incredible Machine, or its puzzles

3. **Off-task conversation** – talking about any other subject
4. **Off-task solitary behavior** – any behavior that did not involve The Incredible Machine or another individual (such as reading a magazine or surfing the web)
5. **Inactivity –** instead of interacting with other students or the software, the student stares into space or puts his/her head down on the desk.
6. **Gaming the System** – sustained and/or systematic guessing, such arranging objects haphazardly or trying an object in every conceivable place. Also, repeatedly and rapidly requesting help in order to iterate to a solution.

The affective categories coded were drawn from [11]. Since many behaviors can correspond to an emotion, the observers looked for students' gestures, verbalizations, and other types of expressions rather than attempting to explicitly define each category. The categories coded were:

1. **Boredom** – behaviors such as slouching, and resting the chin on his/her palm; statements such as "Can we do something else?" and "This is boring!"

2. **Confusion –** behaviors such as scratching his/her head, repeatedly looking at the same interface elements; statements such as "I don't understand?" and "Why didn't it work?"

3. **Delight** – behaviors such as clapping hands or laughing with pleasure; statements such as "Yes!" or "I got it!"

4. **Surprise** – behaviors such as jerking back suddenly or gasping; statements such as "Huh?" or "Oh, no!"

5. **Frustration –** behaviors such as banging on the keyboard or pulling at his/her hair; statements such as "This is annoying!" or "What's going on?!?"

6. **Flow** – complete immersion and focus upon the system [cf. 10]; behaviors such as leaning towards the computer or mouthing solutions to him/herself while solving a problem

7. The **Neutral** state, which was coded when the student did not appear to be displaying any of the affective states above, or the student's affect could not be determined for certain.

Some of these affective categories may not be mutually exclusive (such as frustration and confusion), though others clearly are (delight and frustration). For tractability, however, the observers only coded one affective state per observation.

Past research has suggested that brief observations can be reliable indicators of a student's affective state, whether carried out live [11] or by watching screen-capture videos [12]. 706 observations were collected, for an average of 19.6 observations per student. Inter-rater reliability was acceptably high across all observations – Cohen's [7] κ=0.71 for usage observations, κ=0.63 for observations of affective state.

3. Results

3.1 Overall Results

The two most common behavioral categories observed were working on-task with the software (80% of observations), and talking on-task (9% of observations). The combined frequency of the on-task categories was higher than is common in traditional classrooms [13,14] or intelligent-tutor classrooms [4], though lower than the frequency seen among students playing non-educational action games [17].

Gaming the system was the third most common category of behavior, observed 8% of the time. Although the overall frequency of gaming was higher than in previous observational studies [cf. 4], the percentage of students ever seen gaming – 36% – was within the range observed in previous studies. Both off-task conversation and off-task solitary behavior were quite rare, occurring 0.5% and 0.3% of the time – this frequency is considerably lower than the frequency of off-task behavior in intelligent-tutor classrooms [4] but comparable to the frequency of off-task behavior among students playing non-educational action games [17]. Therefore, one potential interpretation of this finding is that students are generally less likely to go off-task when using environments designed with fun as a primary goal. Another possibility is that our observation technique inhibited students' willingness to be visibly off-task, although this hypothesis does not explain why students engaged in relatively high amounts of gaming the system. And, of course, it is also possible that differences in the population studied (such as age, culture, and type of school) from populations in earlier studies may explain the rarity of off-task behavior in our sample.

The most common affective state observed was flow, coded in 61% of the observations. The dominance of the flow state is similar to results seen in prior studies of affect in students using intelligent tutoring systems [9]. The second most common category was confusion, observed 11% of the time. Boredom (7%), frustration (7%), delight (6%), and the neutral state (5%) were each seen in a small but definite proportion of the observations. Boredom was relatively less common than in previous work studying affect in intelligent tutoring systems (Craig et al, 2004), but frustration was relatively more common. Surprise was the rarest category, but was still observed (3%).

3.2 Affect and the choice to game the system

In this section, we will discuss the relationship between a student's affective state and behavior. We focus on gaming behavior, because gaming the system is known to be associated with poorer learning [4,5,16] and also because gaming was observed with reasonably high frequency in our observations, making statistical inference feasible. We study the relationship between gaming behavior and affective states in two fashions. First, we will study what affective state a student experiences at the time he or she is gaming the system. Second, we will study the antecedents of gaming by investigating what affective states students experience one minute before gaming.

At the exact time a student is gaming, he or she is most frequently bored (39% of the time). Since students are generally bored 7% of the time, they are over 5 times more likely to be bored when they are gaming, a significant difference, $\chi^2(1, N=706) =$

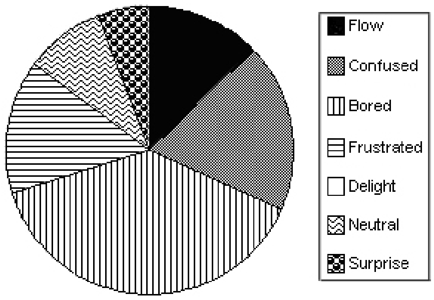

Figure 2. The frequency of affective categories, at the time a student is gaming the system.

95.02, p<0.001. The second most common affective state among a student as he or she games is confusion (19% of the time). Students are generally confused 11% of the time; the difference in frequency between gamers and non-gamers is marginally significant, $\chi^2(1, N=706)=3.01$, $p=0.08$. The third most common affective state among a student as he or she games is frustration (15% of the time). Since students are generally frustrated 7% of the time, they are more than twice as frequently frustrated when they are gaming, a significant difference, $\chi^2(1, N=706)=5.93$, $p=0.01$. Flow, observed only 13% of the time among gaming students (61% of the time overall), is much less common when a student is gaming, $\chi^2(1,N=706)=57.68$, $p<0.001$. The neutral state (9%) and surprise (6%) were not significantly related to gaming. Delight and gaming were never seen in conjunction, in any observation. The overall frequency of each affective state at the time of gaming is shown in Figure 2.

Another way to analyze the relationship between affective states and gaming is to investigate which affective states serve as antecedents to gaming. We can determine what affective states are antecedents to gaming by looking at the probability of gaming one minute after an affective state is observed.

According to our data, if a student is bored, he or she games one minute later 21% of the time, over twice the overall average frequency of gaming, $\chi^2(1, N=634)=10.45$, $p<0.01$. If a student is confused, he or she games one minute later 17% of the time, twice the overall average frequency of gaming, $\chi^2(1, N=634)=7.73$, $p<0.01$.

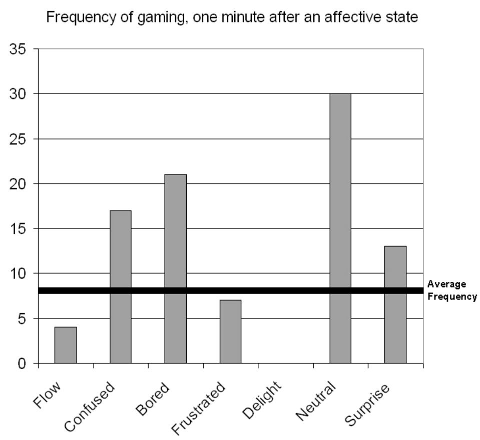

Figure 3. The frequency of gaming one minute after being in a specific affective state.

In addition, if a student is neutral, he or she games one minute later 30% of the time, almost four times more than the overall average, $\chi^2(1, N=634)=22.58$, $p<0.001$.

However, if a student is frustrated, he or she games one minute later only 7% of the time, not significantly different than the overall average frequency of gaming, $\chi^2(1, N=634)=0.05$, $p=0.83$. If a student is surprised, he or she games one minute later 13% of the time, also not significantly different than the overall average frequency of gaming, $\chi^2(1, N=634)=0.40$, $p=0.53$.

If a student is in flow, he or she games one minute later only 2% of the time, much lower than the overall average frequency of gaming, $\chi^2(1, N=634)=22.89$, $p<0.001$. Finally, if a student is delighted, he or she never games one minute later, within our data. The frequency of gaming behavior, one minute after each affective state, is shown in Figure 3.

Hence, boredom and confusion both co-occur with gaming, and are antecedents to gaming the system. Curiously, however, frustration does not appear to be an antecedent to gaming, although it co-occurs with gaming. Additionally, delight and flow are negatively associated with gaming the system, both at the same time and one minute later.

Unexpectedly, the neutral affective state also appears to be an antecedent to gaming. This somewhat surprising finding may suggest that an affective state we did not include in our coding scheme may also be related to gaming (and that this state was coded as neutral by the observers, for lack of a better way to code it.)

4. Discussion and Conclusions

In this study, we have studied student usage choices and affective states in a simulation problem solving environment, The Incredible Machine. One finding is that off-task behavior is quite rare in The Incredible Machine, as in action games [17] but in contrast to intelligent tutoring systems [4]. At the same time, gaming the system is at least as common in The Incredible Machine as it is in intelligent tutors. A potentially interesting question for future work is whether students view gaming the system in the same fashion within game-like simulation learning environments as in intelligent tutoring systems. Students appear to know that gaming the system is inappropriate within tutoring systems, since they hide their gaming behavior from their teachers. Do they believe that gaming the system is inappropriate behavior within more game-like environments? Gaming the system is often not considered appropriate behavior within games (as shown by the lack of respect some game-players have for over-use of cheats and hints within games), but there is also the sense that since games are primarily for fun, it is acceptable to use them in any fashion. If students transfer a sense that gaming the system is appropriate behavior from pure-entertainment games to educational games, there may be implications for how educational games should respond when students game the system.

In terms of the relationship between gaming the system and affective states, we find that boredom and confusion both co-occur with gaming behavior and serve as antecedents to it. Frustration co-occurs with gaming, but does not appear to be an antecedent to gaming. This raises some questions about how frustration and gaming are related – does frustration lead to gaming, but too rapidly for observations spaced a minute apart to detect it? Or do students become frustrated when they choose to game the system and still do not immediately obtain the correct answer? Another affective state, the neutral state, also appears to be a strong antecedent to gaming – we currently have no definitive explanation why. In both cases further research, through more fine-grained video analysis or interviews, may help us understand these patterns better.

In general, knowing that boredom and confusion are antecedents to gaming the system suggests that detecting boredom and confusion [cf. 11] might signal a new way for adaptive systems to respond to gaming behavior. A system that knew a student was bored or confused, and that the student had a history of gaming behavior, could offer the type of supplementary support shown to help gaming students learn better [cf. 3] before the student even starts to consider gaming the system.

Acknowledgements

We would like to thank Anita Rilloraza and the Kostka School of Quezon City for assistance in conducting the studies at the schools, and Genaro Rebolledo-Mendez and Ulises Xolocotzin for useful suggestions and feedback. This work was funded in part by a fellowship from the Learning Sciences Research Institute.

References

[1] Arroyo, I., Woolf, B. (2005) Inferring learning and attitudes from a Bayesian Network of log file data. *Proceedings of the 12th International Conference on Artificial Intelligence and Education*, 33-40.

[2] Baker, R.S.J.d. (in press) Modeling and Understanding Students' Off-Task Behavior in Intelligent Tutoring Systems. To appear in *Proceedings of ACM SIGCHI: Computer-Human Interaction 2007.*

[3] Baker, R.S.J.d., Corbett, A.T., Koedinger, K.R., Evenson, E., Roll, I., Wagner, A.Z., Naim, M., Raspat, J., Baker, D.J., Beck, J. (2006) Adapting to When Students Game an Intelligent Tutoring System. *Proceedings of the 8th International Conference on Intelligent Tutoring Systems*, 392-401.

[4] Baker, R.S., Corbett, A.T., Koedinger, K.R., and Wagner, A.Z. (2004) Off-Task Behavior in the Cognitive Tutor Classroom: When Students "Game the System". Proceedings of ACM CHI 2004: Computer-Human Interaction, 383-390.

[5] Baker, R.S., Roll, I., Corbett, A.T., Koedinger, K.R. (2005) Do Performance Goals Lead Students to Game the System? *Proceedings of the International Conference on Artificial Intelligence and Education (AIED2005)*, 57-64.

[6] Beal, C.R., Qu, L., Lee, H. (2006) Classifying learner engagement through integration of multiple data sources. *Paper presented at the 21st National Conference on Artificial Intelligence.*

[7] Cohen, J. (1960) A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement,* 20, 37-46.

[8] Conati, C., Maclaren, H. (2004) Evaluating a probabilistic model of student affect. *Proceedings of the 7th International Conference on Intelligent Tutoring Systems,* 55-66.

[9] Craig, S.D., Graesser, A.C., Sullins, J., Gholson, B. (2004) Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media,* 29 (3), 241-250.

[10] Csikszentmihalyi, M. (1990) *Flow: The Psychology of Optimal Experience.* New York: Harper and Row.

[11] D'Mello, S. K., Craig, S. D., Witherspoon, A., McDaniel, B., Graesser, A. (2005). Integrating affect sensors in an intelligent tutoring system. In "Affective Interactions: The Computer in the Affective Loop Workshop" In conjunction with *International conference on Intelligent User Interfaces,* 7-13.

[12] de Vicente, A., Pain, H. (2002) Informing the detection of the students' motivational state: an empirical study. *Proceedings of the 6th International Conference on Intelligent Tutoring Systems*, 933-943.

[13] Lee, S.W., Kelly, K.E., Nyre, J.E. Preliminary Report on the Relation of Students' On-Task Behavior With Completion of School Work. *Psychological Reports,* 84 (1999), 267-272.

[14] Lloyd, J.W., Loper, A.B. Measurement and Evaluation of Task-Related Learning Behavior: Attention to Task and Metacognition. *School Psychology Review,* 15 (3) (1986), 336-345.

[15] Sierra Online, Inc. (2001) *The Incredible Machine: Even More Contraptions.*

[16] Walonoski, J.A., Heffernan, N.T. (2006) Detection and Analysis of Off-Task Gaming Behavior in Intelligent Tutoring Systems. *Proceedings of the 8th International Conference on Intelligent Tutoring Systems,* 382-391.

[17] Ziemek, T. (2006) Two-D or not Two-D: gender implications of visual cognition in electronic games. *Proceedings of the 2006 Symposium on Interactive 3D graphics and games,*183-190.