Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

An Examination of Affect and its Relationship with Learning among Students using SimStudent

Michelle BANAWAN[1,2], Maureen VILLAMOR[1,3],Yancy Vance PAREDES[1,2], Cesar TECSON[1,4], Wilfredo BADOY[1], John Roy GERALDE[1,2], Ma. Mercedes RODRIGO[1]
[1]*Ateneo de Manila University, Philippines*
[2]*Ateneo de Davao University, Davao City, Philippines*
[3]*University of Southeastern Philippines, Davao City, Philippines*
[4]*Cor Jesu College, Digos City, Philippines*
mpbanawan@addu.edu.ph, maui.m78@gmail.com, yvmparedes@addu.edu.ph, catecson@cjc.edu.ph, wbadoy@yahoo.com, jrageralde@addu.edu.ph, mrodrigo@ateneo.edu

Abstract: The goal of this paper was to examine affect-related factors and its relationship with student learning while tutoring an agent called SimStudent. These affect-related factors are the negativity of student self-explanations, the incidence and persistence of student affective states. Secondary school students who were part of this study were asked to teach their SimStudents solve algebra equations and make them pass all the quizzes. Results revealed that students failed to learn which led us to investigate other factors that could have attributed to this failure. Although the non-negatively valenced self-explanations did not have significant relationships with the students' learning gains, the self-explanations were helpful in terms of mathematical content and they generally exhibited positive attitudes when giving the self-explanations. Students also tended to perform better with higher levels of good confusion. Higher levels of boredom were associated with poorer learning. Boredom and confusion were the most persistent but both did not have significant relationships with student learning. Though the negative correlations of the negative self-explanations, incidence and persistence of boredom vis-à-vis learning were not significant, the findings imply that negativity is linked to students' poor performance.

Keywords: Affect, SimStudent, learning by teaching, self-explanation

1. Introduction

SimStudent is an algebra-solving teachable agent that learns problem-solving skills from examples (Matsuda et al., 2011; Carlson, Matsuda, Koedinger & Rose, 2012). A student, who acts as the tutor, gives a problem in a form of mathematical equations with variables which a SimStudent tries to solve one step a time (Matsuda et al., 2012c). Researchers have used SimStudent to investigate learning and the factors that affect learning, e.g. competition (Matsuda et al., 2013), meta-cognitive help (Matsuda et al., 2014), deep vs. shallow learning (Matsuda et al., 2012c), quality of self-explanations (Matsuda et al., 2012a) and others. This paper investigates the relationship between affect and learning among students using SimStudent.

Affect refers to a positive or negative mental state coupled with some combination of physiological arousal, cognitive evaluation, and behavioral expression (Picard, 2000). Affect is interesting because its role is encompassing; be it in decision-making, in perception, in human interaction, or in human intelligence (Picard, 2000). There is an interplay between affect and learning and this claim is supported by recent reports made by affective neuroscience and psychology which suggest that human affect and emotional experience are important as they can influence how humans learn (Ahn & Picard, 2005).

In this paper, we narrate a SimStudent deployment that failed to help students learn. We then examine the relationship between student achievement and the following affect-related factors:*negativity of student self-explanations*, *incidence of student affective states*, and *persistence of student affective states*.

2. SimStudent

SimStudent is a teachable agent that helps students learn linear equation problem-solving skills by teaching (Matsuda et al., 2011). It has been tested and redesigned several times, resulting in insights regarding the effects of learning by teaching and related cognitive theories to explain when and how students learn by teaching (Matsuda et al., 2012a; Matsuda et al., 2012b; Matsuda et al., 2013).

SimStudent is a synthetic pedagogical agent that acts as a peer learner. It learns procedural skills from examples. SimStudent attempts to solve a problem given by the student one step at a time. If SimStudent cannot perform a step correctly, it asks the student for a hint. To respond to this request, the student has to demonstrate the step.

This study used the self-explanation version of SimStudent, where the SimStudents ask their tutors to provide explanations for their tutoring decisions, e.g. "*Why should I do this problem?*" or "*But I tried that move earlier. Why doesn't it work now?*" Students could choose a response from a drop-down list or create freeform responses. SimStudents do not understand these self-explanations. It was included in this version to see the effect of self-explanation for tutor learning.

3. Methods

3.1 Participants

The study took place in one high school in Davao City, Philippines. Two (2) first year high school (Grade 7) sections with an average of 36 students per class were enlisted in this study. All students were taking an algebra class. There were 72 study participants in all with ages ranging from 12 to 14. The average age of the participants was 13.5 years old.

3.2 Structure of the study

The actual experiment was comprised of one session designed for the SimStudent orientation and pre-test, three 60-minute sessions for the actual use of SimStudent in the computer laboratory, and one session for post-test and debriefing.

When students used the software, they tutored their SimStudents in solving equations with variables on both sides. They were informed that their goal was to help their SimStudents pass all four (4) sections of the quiz.

3.3 Measures

There were three types of data collected during the experiment: written test data, system logs, and human observations. Students took pre- and post-test before and after the intervention. Three versions of isomorphic tests, tests A, B, and C, were randomly used to counterbalance the test differences in the pre- and post-tests. The test was divided into five parts. Parts 1, 3, and 5 constituted procedural knowledge while parts 2 and 4 constituted conceptual knowledge. We only used the procedural test scores as the learning outcome measure for the current analysis.

The system automatically logged all of the participants' activities including problems tutored, feedback provided, steps performed, examples reviewed, hints requested, and quiz attempts.

Finally, human observers noted students' affect and behavior as they used SimStudent. Observers followed the Baker-Rodrigo Observation Method Protocol (BROMP) 1.0 (Ocumpaugh, Baker & Rodrigo, 2012). The behaviors of interest were *On Task*, *Giving/Receiving Answers*, *Other On Task Conversation*, *Off Task Conversation*, *Off Task Solitary*, *Inactive* and *Gaming* while the affective states were *Boredom*, *Confusion*, *Delight*, *Surprise*, *Flow*, *Frustration*, and *Neutral*.

$$\text{Normalized gain} \quad \frac{\%\text{Post-test score} - \%\text{Pre-test score}}{1 - \%\text{Pre-test score}} \qquad (1)$$

4. Student achievement

Of the 72 participants, only 50 had complete test scores and log data. The analysis that follows is limited to this subset.

As already mentioned earlier, this deployment of SimStudent failed to help students learn. The mean scores of the procedural skill test and their standard deviations are shown in Table 1. The result of the pre-test and the post-test showed that the students did not learn from SimStudent. The average normalized gain was -0.18. Students did not have any classroom instructions between the pre-test and post-test.

Table 1. Mean Test Scores ± SD for Pre-test and Post-Test

Pre-test (%)	Post-test (%)
37.63 ± 10.67	35.58 ± 11.38

With p=0.42, the difference between the pre-test and post-test of the students was not statistically significant.

5. Affect-related factors and their relationships with student learning

In this section, we examine several affect-related factors and their relationship with learning. These include the poor quality of student self-explanations, negative attitudes towards SimStudent, and negative affective states.

5.1 Valence of Self-explanations

There is evidence that self-constructed explanations could be an effective way to facilitate learning because the students are able to reflect in the process when they work out the examples themselves based on prior knowledge (Chi, 1996; Matsuda et al., 2012a). Occasionally, SimStudent asked the students what, how, and why questions. Students made their responses by typing their answers or instructions on the given placeholder.

Students' self-explanation responses to SimStudent were classified into 8 categories using a Self-Explanation Coding Manual by Carnegie Mellon University. Eight (8) (N1, N2,.., N8) different codes were used to classify the type of explanation. Table 2 shows the coding scheme along with the summary of the results of the self-explanations.

Table 2. Coding Scheme for Classifying Self-Explanations

Code	Description	Total	%
N1	The input must include a math concept-oriented explanation why the student entered the problem, why SimStudent's performance was wrong, or why the student did a particular demonstration.	53	11.57
N2	The input only provides a math-related explanation of how to solve the problem.	259	56.55
N3	The input blames SimStudent for an incorrect action on the current problem solving process.	10	2.18
N4	The input is related to math but is vague and abstract. It does not include a math-concept-oriented justification for the student's action.	26	5.68
N5	The input is an admission on the part of the student that he/she made a mistake.	1	0.22
N6	The input is an admission on the part of the student that he/she does not know the answer to SimStudent's question.	5	1.09
N7	The input does not address SimStudent's question or the input is just a number.	104	22.71
N8	The input does not fit into the other categories.	0	0

Code	Description	Total	%
Total		458	100

Two coders independently labeled 651self-explanations based on the coding scheme. Inter-rater reliability was acceptable with Cohen's Kappa=0.66. The 651 self-explanations were produced by 50 students who completed the pre-test and post-test. Four hundred fifty-eight (458) self-explanations were retained after discarding the disagreements between coders.

We computed for the percentage of helpful self-explanations per student using the following formula:

$$(N1+N2)/ \text{TotalLabels} \qquad (2)$$

Helpful self-explanations are students' explanations that were found appropriate in terms of mathematical content.

The percentages of helpful self-explanations were correlated with the individual learning gain of the students. We found that there was no correlation between the helpfulness of self-explanations and learning ($r=0.04$; $p=0.94$)

Was there perhaps a link between the emotional valence of the self-explanations and learning? In other words, could students' positive or negative attitudes towards SimStudent be related to the students' skills in math? Working independently, two coders labeled the 651 self-explanations for valence. A self-explanation was labeled as negative if it manifested negative emotions like anger, fear, and frustration (e.g. shouting at the tutee by typing in all caps like "*I DON'T KNOW HOW, I'M NOT GOOD IN MATH*", using exclamatory marks LIKE "*multiply 3!!!*", not answering, giving irrelevant answers like "*loser*", and the like). A self-explanation was labeled non-negative if it was polite, patient or neutral (e.g. "*You shall subtract 2 in both sides*", "*Yes, transfer 6 to the right side*", "*yes*", "*no*"). Inter-rater reliability was acceptably high, with Cohen's Kappa=0.78

After discarding the disagreement between coders, 554 self-explanations labeled for valence remained. Sixteen percent (16%) (89 out of 554) of the total self-explanations had negative valence. On the other hand, eighty-four (84%) (465 out of 554) had positive valence.

The percentages of self-explanations labeled as negative valence were correlated with the learning gains of the students. We found a low negative correlation between negative self-explanations and learning ($r=-0.27$) however the correlation was not significant ($p=0.56$).

5.2 The incidence of boredom and confusion

Is it possible that poor learning gains were related to the incidence and persistence of confusion or boredom? Both confusion and boredom are interesting as previous studies have shown their relationship to learning (Baker, D'Mello, Rodrigo & Graesser, 2010; D'Mello & Graesser, 2012). Confusion has a positive and negative dimension. It has been found that positive or good confusion can motivate learners to exert more effort to learn while negative or bad confusion can cause learners to give up and disengage from a learning task (D'Mello & Graesser, 2012). Boredom has been found to lead to poor learning gains (Rodrigo et al., 2009; Baker, D'Mello, Rodrigo & Graesser, 2010) and non-productive learner behaviors like gaming the system (Baker, D'Mello, Rodrigo & Graesser, 2010).

As mentioned earlier, 72 students participated in this experiment. For every ten of these students, one pair of affect observers used the Baker-Rodrigo Observation Protocol Method (BROMP) to note the affective states of the students. Unfortunately, only one pair of coders had an acceptable inter-rater agreement, with Cohen's Kappa=0.77. Data from three students was excluded because they had no pre- or post-tests. Hence, for this section, data from only 7 students was included.

With a total of 78 observations, the average for each of the affective state per student was computed to determine the incidence of the affective states. This was further averaged across all seven (7) students and the computed average boredom affect was 13.31% (SD=0.06) while the average confusion affect was 6.57% (SD=0.04).

Learning gains and boredom had a moderate negative correlation (r=–0.61) but this relationship was not significant (p=0.14). Similarly, learning gains and confusion were moderately correlated (r=0.54) but the relationship was not significant (p=0.21). Although not statistically significant, the directionality of the correlations was interesting. Students who tend to be confused tend to do better on the post-test, while students who are bored tend to do worse.

The positive correlation of confusion and learning gains implies that students may have experienced a good form of confusion.

5.3 Persistence of boredom and confusion

Finally, we tried to determine whether the persistence of affective states was related at all to learning gains. We computed each affective state's transition likelihood metric ***L*** (D'Mello, et al., 2005), which is statistically equivalent to Cohen's Kappa.

L is computed as follows:

$$L = \frac{\Pr(NEXT \mid PREV) - \Pr(NEXT)}{(1 - \Pr(Next))} \qquad (4)$$

The ***L*** value for each student was computed for a given transition, followed by the mean and the standard error across students. Using the two-tailed test for one sample, we can then determine if a given transition is significantly more likely than chance (chance = 0), given the base frequency of the next state.

Both boredom-boredom and confusion-confusion transitions were not significant (Table 4).

Table 4. L-Values and SD for Confusion and Boredom

Affective State	L-Value	Standard Deviation	p
Boredom	0.11	0.22	0.34
Confusion	0.02	0.14	0.99

The ***L*** values of boredom and confusion were correlated with the learning gains of the students. The results showed that there was a moderate negative correlation between boredom and the learning gains however the relationship was not significant (r=-0.46, p=0.28). The relationship between confusion-confusion and learning gains was practically zero and not significant (r =-0.01 p=0.98).

6. Summary of Contributions, Limitations, and Future Work

The goal of this paper was to investigate the relationship between student learning gains and affect-related factors. Since student pre-test and post-test scores were not significantly different, we attempted to examine some of the factors that may be related to students' failure to learn. We first examined the quality of students' self-explanations. Majority of the students' self-explanations were helpful in terms of mathematical content. Sixty-nine (69%) of the total self-explanations deliberately instructed SimStudent how to solve the algebra equations. However, the percentages of helpful self-explanations were not significantly correlated with the students' learning gains.

We then examined the valence of the self-explanations. The large majority of self-explanations (84%) were non-negative, i.e. students generally gave helpful, constructive advice. The percentages of negatively valenced self-explanations and learning gains were negatively correlated but the relationship was not significant.

We examined whether certain affective states were related to students' lack of achievement. Although the correlations were not significant, student confusion tended to be correlated with performance while boredom tended to be correlated with poorer learning.

Finally, we examined first whether certain affective states persisted and whether there was a relationship between the student achievement and the persistence of the boredom and confusion. We found that boredom and confusion tended to persist, although the persistence of these affective states was not significant. There were also no significant relationships between persistent boredom and learning gains as well as persistent confusion and learning gains.

Though none of the correlations were significant, the overall trend of the findings imply that negativity goes with poor performance. Negative self-explanations, incidence of boredom, and the persistence of boredom were negatively correlated with the students' learning gains.

For future work, additional, reliable human observations would be very helpful.

Acknowledgements

The authors would like to thank Noboru Matsuda for his research contributions, particularly on SimStudent, that guided us in the completion of this paper.

References

Ahn, H. & Picard, R. W. (2005). Affective-Cognitive Learning and Decision Making: A Motivational Reward Framework for Affective Agents. *Affective Computing and Intelligent Interaction Lecture Notes in Computer Science*, 3784, 866-87.

Chi, M. T. (1996). Constructing self-explanations and scaffolded explanations in Tutoring. *Applied Cognitive Psychology*, *10* (Spec Issue), S33-S49.

D'Mello, S. K., Craig, S. D., Witherspoon, A., McDaniel, B., & Graesser, A. (2005). Integrating Affect Sensor in an Intelligent Tutoring System. *Proceedings of Affective Interactions: The Computer in the Affective Loop Workshop* (pp. 7-13).

D'Mello, S. K., & Graesser, A. (2012). Dynamics of Affective States During Complex Learning. *Learning and Instruction*, *22* (2), 145-157.

Matsuda, N., Cohen, W. W., Koedinger, K. R., Keiser, V., Raizada, R., Yarzebinski, E., et al. (2012a). Studying the Effect of Tutor Learning Using a Teachable Agent that asks the Student Tutor for Explanations. In M. Sugimoto, V. Aleven, Y. S. Chee & B. F. Manjon (Eds.), *Proceedings of the International Conference on Digital Game and Intelligent Toy Enhanced Learning* (DIGITEL 2012) (pp. 25-32). Los Alamitos, CA: IEEE Computer Society.

Matsuda, N., Cohen, W. W., Sewall, J., Lacerda, G., &Koedinger, K. R. (2007). Predicting students performance with SimStudent that learns cognitive skills from observation. In R. Luckin, K. R. Koedinger& J. Greer (Eds.), *Proceedings of the International Conference on Artificial Intelligence Education* (pp. 467-476). Amsterdam, Netherlands: IOS Press.

Matsuda, N., Griger, C. L., Barbalios, N., Stylianides, G., Cohen, W. W., &Koedinger, K. R. (2014). Investigating the Effect of Meta-Cognitive Scaffolding for Learning by Teaching. In S. Trausen-Matu& K. Boyer (Eds.), *Proceedings of the International Conference on Intelligent Tutoring Systems*: Springer.

Matsuda, N., Yarzebinski, E., Keiser, V., Raizada, R., Stylianides, G., Cohen, W. W., et al. (2011). Learning by Teaching SimStudent – An Initial Classroom Baseline Study comparing with Cognitive Tutor. In G. Biswas & S. Bull (Eds.), *Proceedings of the International Conference on Artificial Intelligence in Education* (pp. 213-221): Springer.

Matsuda, N., Yarzebinski, E., Keiser, V., Raizada, R., Stylianides, G., &Koedinger, K. R. (2013). Studying the Effect of Competitive Game Show in a Learning by Teaching Environment. *International Journal of Artificial Intelligence in Education*, *23* (1-4), 1-21.

Matsuda, N., Yarzebinski, E., Keiser, V., Raizada, R., William, W. C., Stylianides, G., et al. (2013). Cognitive anatomy of tutor learning: Lessons learned with SimStudent. *Journal of Educational Psychology*, *105*(4), 1152-1163.

Ocumpaugh, J., Baker, R. S.J.d., & Rodrigo, M. M. T. (2012). Baker-Rodrigo Observation Method (BROMP) 1.0 Training Manual version 1.0.

Picard, R. W. (2000). Affective Computing. MIT Press.