Multi-Site Evaluation of SimSE

Emily Navarro and André van der Hoek
University of California, Irvine
Department of Informatics
Irvine, CA 92697-3440 USA
+1 949-824-5960

{emilyo, andre}@ics.uci.edu

ABSTRACT

In this paper, we describe a multi-site evaluation of SimSE, an educational software engineering simulation game. This study was designed to build on our previous experience of evaluating SimSE in courses and controlled lab settings at UC Irvine, in order to validate our findings and discover any factors that come into play when SimSE is used in other institutions. The study consisted of three different universities using SimSE in their respective courses and reporting the results to us. The results confirmed several of our previous findings, as well as highlighted a number of critical considerations that must be taken into account when using SimSE in a course.

Categories and Subject Descriptors

K.3.2 [**Computer and Information Science Education**]: Computer Science Education

General Terms:

Experimentation

Keywords

Software engineering education, educational evaluation, simulation, educational technology, instructional technology, instructional evaluation, educational games

1. INTRODUCTION

For the past eight years, we have been researching, designing, developing, using, and evaluating SimSE, an educational game-based software engineering simulation environment for software process education. Since the completion of our first prototype in 2004, we have been conducting incremental evaluations of SimSE, the results of which have fed back into its continuing development. SimSE has now matured into a stable and usable tool that has been downloaded and used in numerous classrooms throughout the world.

The evaluation presented in this paper was carried out as part of our goal to transition SimSE from a standalone prototype tool to a comprehensive classroom approach. That is, we want to be able to hand instructors not just a tool, but a tool that we know is developed to its full potential; conclusively evaluated with respect to its strengths, weaknesses, and benefits; and accompanied by extensive documentation and suggested instructions for best practices in class use.

Our evaluations of SimSE had been limited to our university (UC Irvine), with the exception of anecdotal comments from others who have used it. Although these experiences allowed us to learn valuable lessons and gather much of the data necessary for developing a comprehensive classroom approach, we knew that formal, monitored evaluations in classrooms at other institutions would eventually be necessary. Such a multi-site evaluation would allow us to examine how students' and instructors' opinions and experiences align with and/or differ from ours.

SIGCSE'09, March 3–7, 2009, Chattanooga, Tennessee, USA.

2. Background

2.1 SimSE

The goal of SimSE is to bridge the gap between the large amount of conceptual software process knowledge given to students in lectures and the comparably small amount of this they actually get to put into practice in the typical class software engineering project. SimSE accomplishes this by allowing students to practice, through a simulator, the activity of managing different kinds of software engineering processes. SimSE is a single-player game in which the player takes on the role of project manager of a team of developers who must successfully complete a "virtual" software engineering project (see Figure 1).

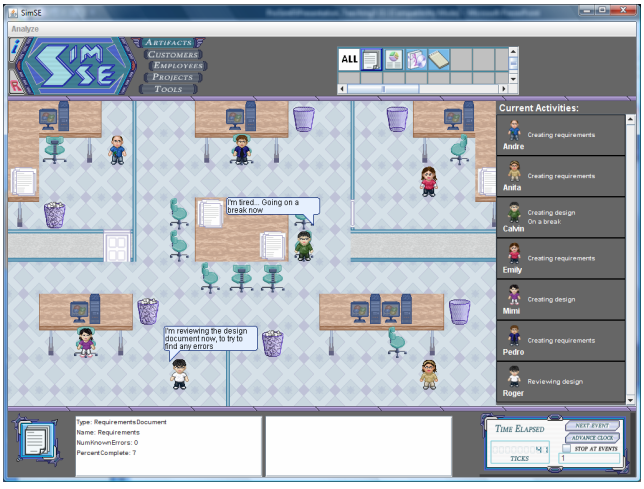

Figure 1: Screenshot of a SimSE Game.

Players use information gleaned from the user interface about their employees, artifacts, customers, tools, and projects to make decisions and take actions, driving the simulation accordingly. At the end of the game, the player receives a score, indicating how well they performed. In addition, an explanatory tool analyzes a player's game, provides them with additional insight into the simulation model, and allows them to run multiple simultaneous branches of a game to explore alternative decisions.

SimSE includes a model builder tool that enables instructors to build models of the particular process(es) they wish to simulate. To date, six SimSE models have been completed: a waterfall model, an incremental model, a code inspection model, a rapid prototyping model, a Rational Unified Process model, and an Extreme Programming model. In each of these, the player is rewarded for following that process's "best practices" and penalized for deviating from them. Full details about SimSE, including its design, game play, and simulation models, can be found in [4], and (shorter) [5].

2.2 Previous Evaluations

Our evaluations at UC Irvine consisted of a series of four studies: (1) An initial pilot study in which 28 students played SimSE and filled out a survey about their experience; (2) Two quarters of using SimSE as an extra-credit assignment in an introductory software engineering course in which students had to play SimSE and answer a set of questions; (3) A pre- and post-test based comparative study of SimSE versus reading a textbook versus lectures; and (4) An in-depth observational study in which we observed students playing SimSE in a one-on-one setting and interviewed them about their experience. Some of our most significant findings were:

- *Students who play SimSE seem to successfully learn the concepts it is designed to teach.*
- *Students find playing SimSE a relatively enjoyable experience.*
- *SimSE has applicability for students of varying abilities and backgrounds.*
- *The learning process of a SimSE player involves the theories of Discovery Learning, Learning through Failure, Constructivism, Learning by Doing, Situated Learning, and Keller's ARCS.*
- *SimSE is most educationally effective when used as a complementary component to other teaching methods.*
- *Providing students with adequate and proper instruction in playing SimSE is critical for its effective use.*
- *Providing students with a set of guiding questions to answer when playing SimSE is critical for its effective use.*

For further details regarding these previous studies and their results, see [6].

3. Multi-Site Study

As part of our plan to expand SimSE into a comprehensive classroom approach, we built on our on-site evaluations by conducting a multi-site study in which three different universities (two large public research universities and a mid-size public historically black university) incorporated SimSE into a course. As mentioned previously, this study was designed to examine how our experiences with using SimSE in courses at UC Irvine align with those in other settings—namely, those in which we are not directly involved or in control.

3.1 Setup

Sites 2 and 3 both used SimSE in software engineering courses, while Site 1 used it in a senior research seminar course (despite our request for it to be used in a software engineering course; regrettably, we did not become aware of this until the study was over). However, 13 of the 14 Site 1 students had previously taken and passed a software engineering course, so they at least had the background knowledge necessary to play the game. Site 1 had 14 students enrolled in the course, Site 2 had 20 students, and Site 3 had 50. All three instructors were compensated with one half-month of summer salary for their participation.

All students in each course were given the option of completing an extra-credit exercise that consisted of playing three SimSE models (rapid prototyping, Rational Unified Process (RUP), and inspection) and answering a set of questions about them. These questions were specifically written in such a way that the students had to play the game in order to find out the answer. We strongly recommended to each instructor that they at least touch on the relevant process models in their lectures and allow the students at least 3 weeks to complete the assignment. Although this may seem excessive, our previous experience with SimSE has shown that such a time commitment is necessary for playing the three models enough to be able to answer the assigned questions. In addition, we gave them a key to use in grading the questions. Aside from those instructions, the instructors were given freedom in the other specifics of the assignment, including how to introduce SimSE and the assignment and how much credit to make the assignment worth.

Upon completion of the courses, we gathered the students' scores on the assignment and their final course grades (both anonymized), and asked both students and instructors to fill out a questionnaire about their experience with SimSE.

3.2 Results

In this subsection, we present the raw results of the multi-site study, which will be analyzed and interpreted in Section 3.3. For the purposes of comparison, we also include the UC Irvine results in these subsections. We report statistically significant results, but omit the details of some statistical tests due to the large number of results and space limitations.

3.2.1 Participation

Nearly all of the students enrolled in each course chose to participate in the assignment—14 out of 14 for Site 1, 19 out of 20 for Site 2, and 48 out of 50 for Site 3. (In the two UC Irvine courses that used SimSE, 12 of 24 students from the first course, and 58 of 87 from the second course participated.)

3.2.2 SimSE Assignment Scores

Site 3 had the highest average score on the assignment (12.75 points out of 15), followed by UC Irvine (12 points), Site 2 (10.2 points), and Site 1 (6.9 points), (see Figure 2). A one-way ANOVA showed that SimSE assignment scores differed significantly across the three sites, $F(3, 151) = 38.89$, $p < .001$. Tukey post-hoc comparisons of the four groups indicate significant differences between all groups, except Site 3 and UC Irvine.

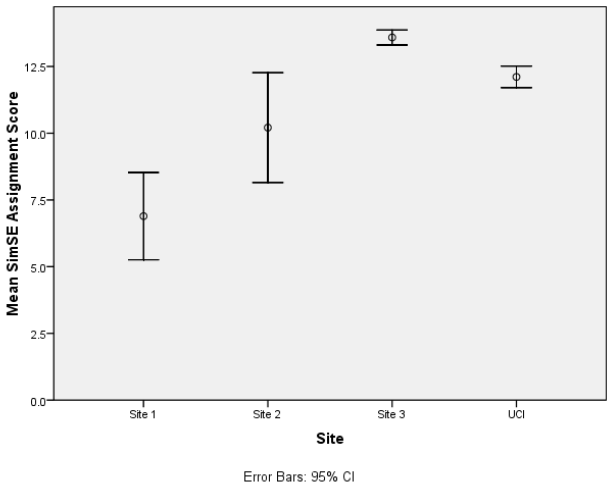

Figure 2: Average Scores on SimSE Assignment by Site (out of 15 Points).

3.2.3 Correlation between Final Course Grades and SimSE Assignment Scores

Testing for a correlation between a student's final course grade (not including points earned on the SimSE assignment) and their score on the SimSE assignment showed no significant correlation for any site except for Site 3 (the site with the highest overall assignment score average). For this site, there was a significant positive correlation, $r(47) = .683$, $p < .001$, (see Figure 3).

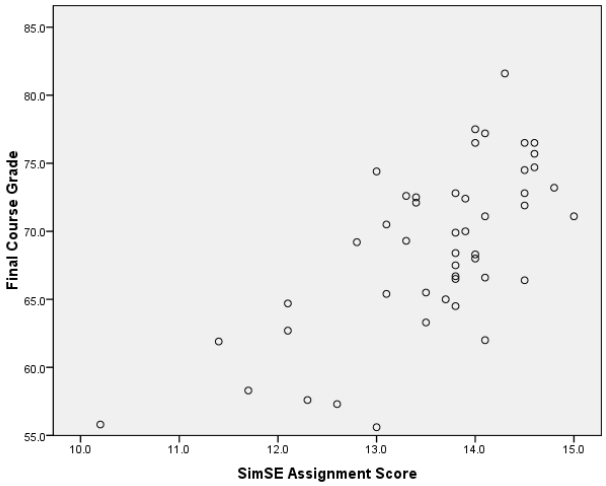

Figure 3: Positive Correlation between Final Course Grade and SimSE Assignment Score for Site 3.

3.2.4 Student Questionnaire Responses

A summary of the differences in student questionnaire responses for the questions that asked for numerical (rather than free-form) answers is shown in Figure 4. (Questions 8, 9, 12, 13, 16, and 17 were only asked in the multi-site usage, not in UC Irvine usage.) All ratings were on a scale of 1 (lowest) to 5 (highest). A one-way ANOVA was used to test for differences among each question for each site. Responses differed significantly across the sites for all of the questions except ease of play. Tukey post-hoc comparisons of the groups indicate that, with one exception, all of these differences resulted from Site 3's significantly higher rankings than the other sites.

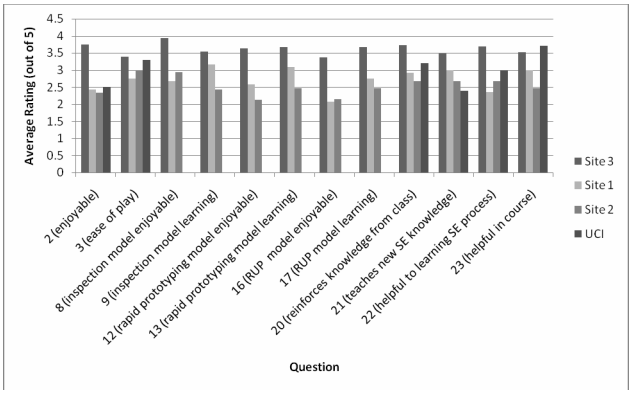

Figure 4: Student Questionnaire Responses Grouped by Site.

In general, Site 2 students had the lowest overall rankings (although not statistically significant), ranking SimSE lowest in 7 out of 12 questions. Site 1 was next lowest (again, not statistically significant), ranking SimSE lowest in 4 out of 12 questions. UC Irvine students' ranked lowest in only one question—how well SimSE teaches new software engineering knowledge.

In addition to comparing questionnaire responses between the different sites, we also compared the responses of off-site students with industrial software engineering experience versus those with no experience (see Figure 5). There were only statistically significant differences in ratings for rapid prototyping model learning, $F(1, 63) = 7.025$, $p = .010$, and for SimSE's helpfulness in the course, $F(1, 109) = 6.766$, $p = .011$, both of which with no experience rated higher.

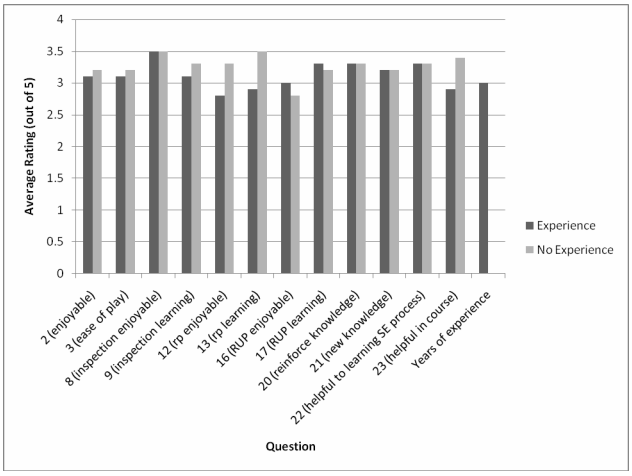

Figure 5: All Off-Site Industrial Experience Differences in Questionnaire Results.

We also compared the questionnaire responses between males and females (see Figure 6) and found that the only statistically significant differences were in ratings for inspection model enjoyability, rapid prototyping model enjoyability, rapid prototyping model learning, and SimSE's ability to teach new software engineering knowledge, all of which females ranked higher than males.

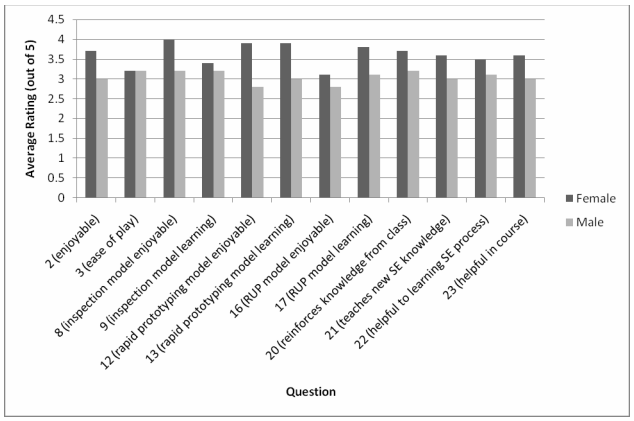

Figure 6: All Off-Site Gender Differences in Questionnaire Results.

3.2.5 Instructor Questionnaire Responses

The instructor questionnaire responses allowed us to understand the similarities and differences in how each instructor incorporated SimSE into their course, so as to give us a framework for interpreting the similarities and differences in the student results. The major varying factors in each site's approach were:

- *Instructor class time spent on SimSE activities.* This varied from 30 minutes (for Site 1) to 2 hours (for Sites 2 and 3). (UC Irvine used approximately 10 minutes of class time.)
- *Instructor activities spent on SimSE.* The instructor from Site 3 had the most extensive in-class activities, including: (1) an in-depth demo of a semi-complete SimSE game, (2) an explanation of SimSE's background, its simulation models, and project management in general, and (3) a presentation by a student (who had already played SimSE) about important aspects of running the game and how it compares to real-world experiences. The instructors from the other sites (including UC Irvine) only presented a few-minute demo SimSE's interface, and gave an introduction to the relevant process models.
- *Percentage of the final grade the SimSE assignment was worth.* The Site 3 instructor (as well as the UC Irvine instructors) made the assignment worth 10% of the final grade, while Sites 1 and 2 made it worth only 5%.
- *Time allowed for completing the assignment (and manner in which it was given).* Students at Site 1 were given the least amount of time to complete the assignment, and were also given it quite differently than the others. They were given 1 week of out-of-class time to play the rapid prototyping model, 1.5 hours of in-class time to play the inspection model, and 3 hours of in-class time to play the RUP model. The questions for each model were given to the students in class after they had finished playing each model, to answer immediately. In contrast, students at the two other sites were given the questions when the assignment was first given, and then allowed several weeks (six for Site 2, eight for Site 3, and five for UC Irvine) to both play the games and answer the questions.

In their questionnaire responses, the instructors also gave several suggestions for ways in which to increase the educational effectiveness of SimSE. All three sites expressed that a more extensive in-class demo, showing either a complete or semi-complete SimSE game, would have helped tremendously in getting students started on the assignment. Other suggestions included allowing students to play in teams, adding a competitive aspect to the exercise, making the assignment mandatory, and assigning an essay in which students compare their experiences playing SimSE to those in their class project.

3.3 Discussion

The multi-site evaluation both served to confirm some of our findings from our previous studies, as well as to bring to light some important considerations that need to be emphasized when using SimSE in a course. Both are listed below.

SimSE has applicability for students of varying abilities and backgrounds. The lack of correlation between SimSE assignment scores and final course grades for Sites 1 and 2 is in line with the UC Irvine results, for which there was also no correlation. It is unclear why there was a correlation for only Site 3. Because three out of the four sites (UC Irvine included) showed no correlation, we can probably conclude that, for most students, there is no correlation between SimSE assignment score and final course grade, suggesting that SimSE is applicable for (and can teach) students of varying academic aptitudes. Furthermore, looking at both the multi-site evaluation and previous UC Irvine usage shows that there are no consistent significant trends in questionnaire rating differences between males and females or experienced and non-experienced students, suggesting that SimSE is also equally applicable to students from various backgrounds.

Students who play SimSE seem to successfully learn the concepts it is designed to teach. Assignment scores at all four sites averaged high, with the exception of Site 1, whose low scores can be explained through significant varying factors in the administration of the assignment (to be discussed later). (We are unsure why Site 2 also had significantly lower scores than UC Irvine and Site 3, but their scores were only moderately low, still averaging in the 2/3 range.) Moreover, examining the scores more closely reveals a phenomenon we also observed in the UC Irvine in-class usage: most low scores can be attributed to the student just not attempting (an) entire model(s). In other words, they only partially completed the assignment, skipping over one or more model's entire set of questions. This was true even for the Site 1 students, who had notably lower scores than the other sites, but also seemed to skip entire models quite frequently.

Students find playing SimSE a relatively enjoyable experience. The ratings of the students in the remote sites were generally comparable to those of the UC Irvine students (mostly in the range of 2.5 to 4), indicating that they liked playing SimSE for the most part, but were not "swept away" by it. This is not surprising, given that the assignment involved real work, real time invested, and real pressure to earn real credit. The participation rate in the assignment was also quite high, indicating that the idea of SimSE interests students (although, to be sure, the extra credit was also part of the draw).

SimSE is most educationally effective when used as a complementary component to other teaching methods. More precisely, when used as part of a course in which relevant knowledge (software processes) is being taught in parallel using other teaching methods (such as lectures and class projects), SimSE is effective at articulating and aiding understanding of this knowledge. This was most evident in the low scores and ratings of Site

1. We suspect that part of the reason for these low numbers was the fact that the course was not a software engineering course, so the students were not learning any of the concepts in parallel (with the exception of 30-minute lectures on two of the process models and two handouts on the other). Therefore, the students probably found it more difficult than those at the other sites who were learning the concepts in parallel.

Providing students with adequate and proper instruction in playing SimSE is critical for a maximally educational experience. This need for adequate instruction was shown repeatedly in our previous evaluations, especially in our observational experiment, in which we were able to see firsthand some of the critical things that players tend to miss if not pointed out and demonstrated explicitly in the instructions [6]. The most obvious manifestation of this in the multi-site study was Site 3's significantly more positive experience than the other sites (higher scores and ratings), along with the Site 3 instructor's significantly more in-depth and extensive introduction of SimSE. Most likely, the extra time and effort expended by this instructor was the most significant contributing factor to this. Thus, instruction must be a carefully planned part of SimSE's use.

Providing students with a set of guiding questions to answer when playing SimSE helps to achieve a maximally educational experience. In our previous comparative and observational studies (in which the questions were not given to the students), we found that these students were less able to discover some of the not so obvious lessons encoded in the simulation models. The Site 1 results corroborate this. We suspect that giving these students the questions as more of an after-the-fact quiz than as a guide in helping the player know what to look for as they play was a major factor in Site 1's low scores and ratings.

An assignment involving SimSE should be made , non-trivial part of the final grade (at least 10%). Students for whom the assignment was worth less than this (5% for Sites 1 and 2) did not do as well on the assignment and did not have as positive perceptions of SimSE than those for whom it was worth 10% (Site 3 and UC Irvine). This suggests that putting more at stake in the assignment provides better motivation for students to put the effort in, learn the concepts, and have a positive experience.

4. RELATED WORK

There have been a handful of other educational game-based software engineering simulations developed, such as SESAM [2], OSS [7], SimVBSE [3], and The Incredible Manager [1]. While these simulations vary in their scope, purpose, and capabilities, the distinction between them and SimSE most relevant to the subject of this paper is the extent to which they have been evaluated, especially in a classroom setting. Evaluations of these simulations have been preliminary and informal in nature (and have not come close to approaching the cross-site angle). In-class usage of these simulations has been minimal and only anecdotally observed and reported on [7]. Others have performed only a single, small out-of-class study each, but have not gone beyond this.

5. SUMMARY AND FUTURE WORK

Our multi-site evaluation of SimSE both confirmed many of our previous findings, and pointed out critical considerations that must be paid attention to when using SimSE in a course. Some of the most critical lessons learned were: (1) SimSE does successfully help students learn software process concepts, (2) providing students with proper instruction in playing SimSE is critical, and (3) providing students with a set of guiding questions to answer while playing SimSE is critical.

To advance our goal of transitioning SimSE from a standalone prototype technology to a comprehensive classroom approach, we are currently compiling all of our results into course modules that are being explicitly designed to support instructors and integrally address classroom methodology. These modules will include such elements as learning objectives, supporting lectures and assignments, instructions for students and instructors, and a video demonstration to use in introducing SimSE to a class.

In the near future, we also plan to explore the other modes of use that have been proposed by both the instructors involved in the multi-site study and others who have used SimSE—such as competitive team-based playing, making SimSE a mandatory course component, and having advanced students build a new SimSE model rather than just play one.

6. ACKNOWLEDGMENTS

We thank the instructors and students at the remote sites for their participation in the study. Effort funded by the National Science Foundation under grant number DUE-0618869.

7. REFERENCES

[1] Dantas, A.R., M.O. Barros, and C.M.L. Werner, *A Simulation-Based Game for Project Management Experiential Learning,* in *2004 International Conference on Software Engineering and Knowledge Engineering.* 2004, P. 19-24.

[2] Drappa, A. and J. Ludewig, *Simulation in Software Engineering Training,* in *22nd International Conference on Software Engineering.* 2000, P. 199-208.

[3] Jain, A. and B. Boehm, *SimVBSE: Developing a Game for Value-Based Software Engineering,* in *19th Conference on Software Engineering Education and Training.* 2006, P. 103-111.

[4] Navarro, E., *SimSE: A Software Engineering Simulation Environment for Software Process Education.* 2006, University of California, Irvine, Doctoral Dissertation: Irvine, CA.

[5] Navarro, E. and A. van der Hoek, *Design and Evaluation of an Educational Software Process Simulation Environment and Associated Model,* in *18th Conference on Software Engineering Education and Training.* 2005, P.25-32.

[6] Navarro, E. and A. van der Hoek, *Comprehensive Evaluation of an Educational Software Engineering Simulation Environment,* in *20th Conference on Software Engineering Education and Training*. 2007, P.195-202.

[7] Sharp, H. and P. Hall, *An Interactive Multimedia Software House Simulation for Postgraduate Software Engineers,* in *22nd International Conference on Software Engineering.* 2000, P. 688-691.