Engaging Game Design Students Using Peer Evaluation

Amber Settle  
DePaul University  
243 S. Wabash Avenue  
Chicago, IL 60604  
(312) 362-5324  
asettle@cdm.depaul.edu

Charles Wilcox  
DePaul University  
243 S. Wabash Avenue  
Chicago, IL 60604  
(312) 362-8638  
wilcox@cdm.depaul.edu

Chad Settle  
University of Tulsa  
800 South Tucker Drive  
Tulsa, OK 74104  
(918) 631-3157  
chad-settle@utulsa.edu

ABSTRACT
Many information technology educators have worked in recent years to develop courses to attract students to the field. As faculty achieve success with technical courses designed to be appealing to a broad audience, it can be hard to maintain the initial excitement particularly as multiple sections of the courses are taught on a continuing basis. In this article we describe a project that added peer evaluation to an assessment in a game design course with a large non-major audience. While controversial, peer evaluation has shown some promise in motivating students to work harder and in improving certain key skills. Consistent with other studies in areas outside of game design, the introduction of peer evaluation showed significant improvement in student engagement.

Categories and Subject Descriptors
K.3.2 [Computers and Education]: Computer and Information Science Education – Curriculum.

General Terms
Measurement, Design, Experimentation.

Keywords
Peer evaluation, game design, non-majors, general education

1. INTRODUCTION
Since the decline in technology-focused enrollments began in 2002, numerous information technology educators have developed approaches to attract more students to the field. Many have seen it as a motivational issue and have worked on designing courses that engage students. Some have had success creating courses that focus on multi-media [12, 13], the Internet [14, 15], or gaming [2, 4, 5, 8, 16, 17, 19, 26]. A number of these new approaches to introductory courses have produced large enrollments for the institutions that created them [12, 19], attracting many non-major enrollments. As faculty achieve success with these courses it can be hard to maintain the initial excitement, particularly as multiple sections of the courses are taught on a continuing basis.

We describe an approach to improving engagement in an introductory game design course with significant non-major enrollment. The course is taught by a relatively small group of faculty and is a challenging course for students with a strong emphasis on writing [19]. Using an approach that others have found productive [9, 10, 11, 20], the course’s game design project was modified to include peer assessment. In particular, students were involved in developing a rubric for the required game design presentations and peer assessment was incorporated into the grade for the projects. This intervention showed significant improvement in student engagement as measured by the course evaluations. We also believe that the project served to improve student perception of the game design project.

1.1 Understanding Peer Evaluation
Peer evaluation is a well-studied and controversial pedagogical tool. There are many definitions of peer evaluation, making it important to agree on the definition being used here. Topping states that peer assessment is “an agreement in which individuals consider the amount, level, value, worth, quality, or success of the products or outcomes of learning of peers of similar status” [23].

There are a number of advantages and disadvantages of peer evaluation discussed in the literature. Peer evaluation can increase student responsibility and autonomy, provide insight into assessment procedures and expectations for high quality work, and provide students with motivation to work harder because they will be evaluated by their peers [7]. Further, peer assessment is advocated as a way of improving employability, particularly on non-technical skills such as oral and written communication [7]. Formative feedback from peers is associated with improvements in student confidence and presentation [23]. On the other hand, there are numerous concerns about using peer evaluation in the assignment of grades. This is due in part to issues with lack of expertise, potential bias, and a lack of formal training [7, 25].

One reason to consider using peer evaluation despite the significant drawbacks is that almost all studies show that students hold favorable views of the process [1, 3]. It has been shown that students feel personally motivated by the peer evaluation process and that it inspires them to improved performance [6]. However, students are often concerned about their ability to assess their peers and the responsibility of such a task [7].

One important component of using peer evaluation is ensuring that it is used appropriately. It has been suggested that peer evaluation is a skill that can be learned and that the failure to include appropriate training in peer evaluation can result in failure of the process [1, 21]. A suggestion seen in the literature is that peer assessment should emphasize the learning aspects associated with the process rather than focusing on the assigning of grades [7, 25].
What is interesting for this work is that the literature emphasizes general skills such as writing, and it has been suggested that more work on identifying peer assessment issues in the area of information technology would be useful [1]. There has been a recent resurgence of interest in peer evaluation, in part because of the focus on using information technology to better facilitate the process [22, 24]. The study discussed here is not concerned with utilizing technology for the peer evaluation process, so the focus remains on more fundamental issues.

2. GAMING CURRICULUM
In this section we briefly describe the content and student population for the game design course and explain how peer evaluation was introduced into the course.

2.1 The Game Design Course
The College of Computing and Digital Media (CDM) at DePaul University is one of the largest information technology institutions in the country. Approximately 1800 graduate students and 1300 undergraduates are enrolled in 16 Bachelors and 18 Masters degree programs, and DePaul was one of the first liberal arts universities in the country to offer a four-year degree in Computer Game Development. The Computer Game Development degree emphasizes the importance of students taking courses relevant to their interests early in the program. For this reason, first-year students are required to take three game design or development courses. Originally, and during the time of this study, GAM 224: Introduction to Game Design was the first required course. It has since been replaced by a major-specific game design class, but GAM 224 continues to be taught as a general education course.

GAM 224 has no prerequisites and does no computer game development. Instead students in the course study the principles of game design and use these principles both to analyze existing games and to develop their own original game ideas. The course uses a game design textbook that studies games using three different approaches: formal aspects (rules), experiential aspects (play), and cultural aspects (culture) [18]. The course has a strong emphasis on writing, with three types of assessments (reaction papers, analysis papers, and game design documents) that require significant amounts of writing [19]. GAM 224 is approved for credit in the Arts and Literature Domain of the Liberal Studies Program.

Since GAM 224 was first approved as a course in the Arts and Literature Domain in Winter 2005, there have been 66 sections of the course offered. During the time of this study, non-major enrollments in GAM 224 ranged between 25% and 75% depending on the quarter. Non-majors taking the course include students in the College of Commerce, the College of Communication, and the College of Liberal Arts and Sciences. Motivating a diverse group of students in a writing-intensive course like GAM 224 is a challenge. In the section that follows we describe an approach used by the second author to engage the diverse audience for the game design course.

2.2 The Introduction of Peer Evaluation
Based on an article discussing efforts at improving student communication skills in a gaming course [10], we began a project on peer evaluation in the game design courses in the Fall 2007. Students in the game design course are required to design a game and make course presentations about their project. Since GAM 224 does not have any programming prerequisites, the games that are designed are generally not implemented. Instead students design board, card, or electronic games on paper, focusing on various aspects of the game design process that are important for the learning goals of the course. The game design project is a group effort, and students have various deliverables throughout the quarter that culminate in a final design document and in-class course presentation about their game.

As a part of the study, the students were asked to develop a peer evaluation form for the game design projects and associated presentations. The authors constructed an initial draft of the requirements by which the game design projects and oral presentations would be evaluated, which was heavily influenced by the rubric used in a previous study [10]. During the Fall 2007 and Winter 2008 there were in-class activities beginning in the middle of the quarter that centered on the development and refinement of the project peer evaluation form. Students began this work prior to starting their game design projects so that they would clearly understand how their projects would be assessed.

The evaluation form developed by the students is given below. Each item was rated on a 5 point-scale with 1 indicating the worst rating and 5 the best rating:

1) Slides
   a) Readability
   b) Layout
   c) Images borrowed or created
   d) Spelling and grammar
   e) Understandability
   f) Writing

2) Speech (presentation effectiveness)

3) Presentation content
   a) Meets all project requirements
   b) Design specifications
   c) Core mechanic
   d) Creativity of the redesign

4) Elements presented
   a) The characters in the new game
   b) The terrain of the new game
   c) The objects used in game play
   d) The relationships between the objects in the new game
   e) Scoring
   f) Controls used in the source and redesigned game
   g) The new game’s status bar (health meter, etc.)
   h) Winning and losing conditions

5) Connection to Rules of Play (text) ideas/terminology

The peer evaluations were incorporated into the grade assigned to the group projects, with students completing a peer evaluation for every other student in the class. Each class decided how much of the grade was determined by peer evaluations and how much of the grade would be determined by the instructor. During both quarters the students allocated 15% of the project grade to peer evaluation and 85% of the project grade to instructor evaluation.

Students were required to complete an evaluation of their team members by ranking the contributions of the of each team member. The instructor used this information to help allocate the team grade. For the most part this information was used to distinguish students who either excelled or who failed to actively contribute to the project. Any student who was identified as contributing the least by all team members earned a lower
The project was designed to take advantage of the positives associated with peer evaluation while minimizing the negatives. Students were involved in the design of the instruments to provide them with insight into the assessment process for the game design project. Because GAM 224 involves a significant amount of work yet has a large non-major population, students sometimes suffer from a lack of motivation. It was our hope that involving students in peer evaluation would raise expectations about the quality of work that should be produced and motivate the students to perform well, as is seen in the literature [6]. We wanted to ensure that they felt positively about the process, so students developed and refined the instrument even before they began work on the project. Students were also allowed to decide how much peer evaluation should count in the grade, hopefully mitigating any problems with grade assignment and emphasizing the learning aspects of the exercise [7, 25].

3. STUDENT ENGAGEMENT

To evaluate the impact that the project had on student engagement and satisfaction with the course, two types of data were gathered: course evaluation data from all of the sections taught by the second author and the student survey data.

3.1 Course Evaluations

To determine if the peer evaluation of group project presentations had any impact on the student perception of and satisfaction with the course, we ran regressions on the course evaluation data for each section taught by the second author.

The second author taught 11 sections of the course between Fall 2005 and Winter 2008. In order to net out changes over time and in order to determine the effect of changes made to the course prior to the first quarter of the 2007-2008 academic year we include a variable for time. We also include a dummy variable for the changes, P (for the dummy variable, Post). Our Ordinary Least Squares regression equation is given by $Q_i = \alpha_0 + \alpha_1t + \alpha_2P + \epsilon_i$, where $Q_i$ is an individual question on the course evaluation, $t$ = time, $P = 0$ if the course was taught prior to the changes being made and $P = 1$ if course was taught after the changes were made. Time is included as a proxy in order to account for systematic and immeasurable changes in teaching and/or student attitudes across time. The variable for time will include changes in the course structure, changes in the professor’s teaching methods, and changes in general student attitudes toward their educational experience across time. Time is measured by quarters taught at DePaul (for the Fall Quarter of 2005-2006, $t = 1$, for the Winter Quarter of 2005-2006, $t = 2$, etc.). The dummy variable, Post, is included in order to account for differences between courses due to the changes the professor made to the course prior to the 2007-2008 academic year. If $\alpha_2$ is statistically different from 0, it will indicate a difference in how students view the same course due to the changes.

Student evaluations are conducted for every DePaul CDM course during every quarter. The evaluations are conducted online using a secure login from the DePaul CDM web site, and no identifying information about the student is associated with the evaluation. Course evaluations are completed during the 8th and 9th weeks of the 10-week quarter, although results are not made available to instructors until after grades are submitted. During the time of the study the evaluations consisted of 27 multiple choice questions and several sections for comments. The multiple choice questions ask the student to rate various aspects of the course and the instructor for the course. The ratings for the five university-wide questions are on a scale from 1 to 5 and the ratings are on a scale from 0 to 10 for the twenty-two college-specific questions. In general, a higher number indicates a greater degree of student satisfaction with the area addressed by the question. For DePaul CDM-specific questions, a zero indicates that the student feels the question is not applicable.

The university questions are as follows:

1. Given your experience at DePaul, rate the instructor’s overall teaching effectiveness.
2. The instructor stimulated interest in the subject.
3. Given your experience at DePaul, rate the overall quality of the course.
4. Overall this course increased my knowledge or skills.
5. I found this course to be: (relating to course challenge)

The DePaul CDM-specific questions are of two types: course-specific questions and instructor-related questions. The course-specific questions are:

1. Was this course well organized?
2. Do you feel the course objectives were accomplished?
3. The amount of work you performed outside of this course was: (answers relate to the time spent)
4. How difficult was the course material?
5. The textbook for this course was
6. The supplementary reading for this course was
7. The assignments for this course was (sic)
8. The grading of homework and exams in this course was
9. What is your overall estimate of this course?
10. How valuable was this course in terms of your technical development?

The instructor-related questions are:

1. How would you characterize the instructors (sic) knowledge of the subject?
2. How would you characterize the instructor’s ability to present and explain the material?
3. Does the instructor motivate student interest in the subject?
4. How well does the instructor relate the course material to other fields?
5. Did the instructor encourage participation from the students?
6. Was the instructor accessible outside of class?
7. What was the instructors (sic) attitude? How did they treat you?
8. How well did the instructor conduct, plan and organize classes?
9. Were the instructors (sic) teaching methods effective?
10. Would you recommend this course to another student?
11. Would you take this instructor for another course?
12. Rate the teaching effectiveness of this instructor as compared to other faculty in this school.

The pooled sample including evaluations from all 11 sections of the course yield a total of 316 observations for each course-related question and instructor-related question and 186 observations for the university-wide evaluation questions. There were fewer observations for the university-wide questions since they were not included on the DePaul CDM course evaluations.
until Fall 2007. After dropping responses of zero for each question, the regression equation is run 27 times, once for each question on the evaluation in order to determine which questions have a statistically significant difference across time and for pre/post-change effects.

As seen in Table 1, the coefficients for the university-wide questions 1, 2, 3, and 4 were significantly different from zero across time, with each falling over time. Similarly, the coefficients for university-wide questions 1, 2, 3, and 4 were significantly different from zero pre- and post-course changes with each evaluation question increasing after the changes were made. In all of the tables presented in this section, coefficient estimates are presented with standard errors in parentheses. Those marked with a single asterisk (*) are statistically significant at the 1% level on a two-tailed test, those marked with two asterisks (**) are statistically significant at the 5% level on a two-tailed test, and those marked with three asterisks (***) are statistically significant at the 10% level of a two-tailed test.

Table 1: Regression results for university-wide questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Time</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q-01</td>
<td>-0.242*** (0.076)</td>
<td>0.398* (0.239)</td>
</tr>
<tr>
<td>Q-02</td>
<td>-0.233*** (0.061)</td>
<td>0.653*** (0.191)</td>
</tr>
<tr>
<td>Q-03</td>
<td>-0.269*** (0.088)</td>
<td>0.612** (0.278)</td>
</tr>
<tr>
<td>Q-04</td>
<td>-0.235** (0.094)</td>
<td>0.584** (0.295)</td>
</tr>
<tr>
<td>Q-05</td>
<td>-0.082 (0.088)</td>
<td>0.151 (0.278)</td>
</tr>
</tbody>
</table>

The university-wide questions were begun in the Fall 2007, which is when the change in the course was made. It is interesting that the values for the questions are decreasing over time. The instructor for the course had a difficult time incorporating the data for the game into the course grades in the Fall 2007, and he may have been less enthusiastic about the course during the Winter 2008. However, the use of the peer evaluation project seems to have inspired more enthusiasm in the students and improved the course evaluations.

As seen in Table 2, the coefficients for course-related questions 2, 5, 6, 7, 8, and 9 were all significantly different from zero across time with each coefficient decreasing over time. The coefficients for course-related questions 5, 6, 7, and 8 were all significantly different from zero pre- and post-course changes with each coefficient increasing after the course changes were made.

As seen in Table 3, the coefficients for instructor-related questions 1, 3, 4, 6, 7, 10, 11, and 12 were all significantly different from zero across time with each decreasing over time. The coefficients for instructor-related questions 1 and 7 were significantly different from zero pre- and post-course changes with each increasing after the course changes were made.

It is often the case that faculty pay the most attention to new courses, ones that are new preparations or that they have not taught frequently. The second author has taught the game design course frequently, nearly every quarter since Fall 2005. This helps to explain why course-related questions having to do with the objectives, the textbook, the supplementary reading, and the

Table 2: Regression results for course-related questions

<table>
<thead>
<tr>
<th>Q</th>
<th>Time</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.094** (0.043)</td>
<td>0.435* (0.230)</td>
</tr>
<tr>
<td>2</td>
<td>-0.091 (0.057)</td>
<td>0.085 (0.308)</td>
</tr>
<tr>
<td>3</td>
<td>-0.158*** (0.054)</td>
<td>0.277 (0.288)</td>
</tr>
<tr>
<td>4</td>
<td>-0.172*** (0.066)</td>
<td>0.214 (0.357)</td>
</tr>
<tr>
<td>5</td>
<td>-0.083 (0.077)</td>
<td>-0.338 (0.414)</td>
</tr>
<tr>
<td>6</td>
<td>-0.186** (0.073)</td>
<td>0.556 (0.397)</td>
</tr>
<tr>
<td>7</td>
<td>-0.183*** (0.060)</td>
<td>0.545* (0.322)</td>
</tr>
<tr>
<td>8</td>
<td>-0.074 (0.065)</td>
<td>0.275 (0.348)</td>
</tr>
<tr>
<td>9</td>
<td>-0.101 (0.073)</td>
<td>0.096 (0.398)</td>
</tr>
<tr>
<td>10</td>
<td>-0.170** (0.084)</td>
<td>0.408 (0.452)</td>
</tr>
<tr>
<td>11</td>
<td>-0.135* (0.081)</td>
<td>0.101 (0.434)</td>
</tr>
<tr>
<td>12</td>
<td>-0.140** (0.066)</td>
<td>0.253 (0.357)</td>
</tr>
</tbody>
</table>

Table 3: Regression results for instructor-related questions

<table>
<thead>
<tr>
<th>Q</th>
<th>Time</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.094** (0.043)</td>
<td>0.435* (0.230)</td>
</tr>
<tr>
<td>2</td>
<td>-0.091 (0.057)</td>
<td>0.085 (0.308)</td>
</tr>
<tr>
<td>3</td>
<td>-0.158*** (0.054)</td>
<td>0.277 (0.288)</td>
</tr>
<tr>
<td>4</td>
<td>-0.172*** (0.066)</td>
<td>0.214 (0.357)</td>
</tr>
<tr>
<td>5</td>
<td>-0.083 (0.077)</td>
<td>-0.338 (0.414)</td>
</tr>
<tr>
<td>6</td>
<td>-0.186** (0.073)</td>
<td>0.556 (0.397)</td>
</tr>
<tr>
<td>7</td>
<td>-0.183*** (0.060)</td>
<td>0.545* (0.322)</td>
</tr>
<tr>
<td>8</td>
<td>-0.074 (0.065)</td>
<td>0.275 (0.348)</td>
</tr>
<tr>
<td>9</td>
<td>-0.101 (0.073)</td>
<td>0.096 (0.398)</td>
</tr>
<tr>
<td>10</td>
<td>-0.170** (0.084)</td>
<td>0.408 (0.452)</td>
</tr>
<tr>
<td>11</td>
<td>-0.135* (0.081)</td>
<td>0.101 (0.434)</td>
</tr>
<tr>
<td>12</td>
<td>-0.140** (0.066)</td>
<td>0.253 (0.357)</td>
</tr>
</tbody>
</table>
assignments for the course, as well as the grading of the homework and exams, and the overall estimate of the course would decline over time. The peer evaluation project led the second instructor to reconsider the entire game design project which is a major piece of the course assessments. We believe this led to improved course evaluations for the textbook, the supplementary reading, the assignments for the course, and the grading of homework and exams. The renewed attention to the game design project improved student perception of the readings and the assessments in the course.

As mentioned before, the second author has been a frequent instructor for the game design course. He is also a dynamic individual, thriving best in new and unfamiliar situations. This is clearly seen in the decreasing numbers for the questions having to do with the instructor’s knowledge of the subject, ability to motivate student interest, ability to relate course material to other fields, accessibility outside of class, and attitude, whether the student would recommend the course to another student, take the instructor for another course, and the relative teaching effectiveness of the instructor. The peer evaluation project served as a challenge for the instructor, which enabled him to improve the students’ perception of him regarding his knowledge of the subject, his motivation of students and his attitude. It appears that students found it more motivating to engage in the peer evaluation project as well as perceiving that the instructor treated them better as a result.

3.2 Surveys

Students in each section of the course were asked to complete a survey reflecting on their experience with group projects and peer evaluation. Survey data was gathered for both Fall 2007 and Winter 2008. Unfortunately, the survey data for Winter 2008 was lost so that we had insufficient data to perform a regression. However, a summary of the survey data provides interesting insight into the students’ experience with group design projects and peer evaluation in the game design course.

The survey consisted of questions about the group project independent of the peer evaluation project, questions about the group presentations, and questions about the peer evaluation process. The questions were rated on a 5-point Likert scale.

The questions about the group project were as follows:
1. Working in groups on the design project for this course improved by learning
2. Having more deadlines would have been useful on the group design project
3. There was sufficient instructor feedback about the status of the group project
4. I would have wanted to have group meetings with the instructor to discuss the status of the group project
5. I felt that each student in my group was graded fairly
6. The group project was valuable for this course

The questions about the group presentations were as follows:
1. Presentations for the group projects was a valuable part of the course
2. I learned a great deal by preparing for the presentation
3. I learned a great deal by listening to other students’ presentations
4. Preparations as a part of a game design team helped me to improve by communication skills
5. Participating in the game design presentation in front of the class helped me to improve my communication skills

The questions about the peer evaluation portion of the course were as follows:
1. I learned a great deal by participating in the evaluation of other students’ presentations
2. The feedback I received from my classmates was valuable
3. Designing the peer evaluation form helped me to produce a better presentation
4. The peer evaluation process was a valuable component of the course

Table 4: Summary of student responses to group-project, group-presentation, and peer-evaluation questions

<table>
<thead>
<tr>
<th>Project</th>
<th>Presentation</th>
<th>Peer evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quest. Avg</td>
<td>Quest. Avg</td>
<td>Quest. Avg</td>
</tr>
<tr>
<td>1</td>
<td>4.01</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2.93</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>3.22</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>3.9</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>4.25</td>
<td></td>
</tr>
</tbody>
</table>

On average, students agreed that working in groups on the design project improved their learning, that there was sufficient instructor feedback about the status of the group project, and that the group project was valuable for the course. They neither agreed nor disagreed that having more deadlines would have been useful on the group design project and that they would have wanted to have group meetings with the instructor to discuss the status of the group project. Overall one can conclude from the average student response that they were satisfied with group projects in the course and felt that group projects contributed positively to their learning.

The averages of the numbers for the group presentation questions indicate responses close to agreement for all questions. While not strong support, it does indicate a relative satisfaction with the game design project presentations. The results are stronger than it might seem given that the game design students in general are not comfortable with public speaking and are usually reluctant to give their presentations.

Students were more neutral about the value of peer evaluation. It may be that the process of preparing for peer evaluation led them to better see the benefit of game design presentations, but this is not clearly demonstrated from the data.

4. CONCLUSION

Adding a peer evaluation component to game design presentations had a significant impact on student evaluations in the game design class. Students were uniformly more satisfied with the instructor of the course, perceiving him to be more knowledgeable and with an improved attitude, after the project was added to the course. The project also changed the students’ perception of the course, improving their perception of the textbook, supplementary reading, the assignments, the grading of the assignments and the exams, and the overall estimate of the course.
The survey data, while taken from only a single quarter, indicates that students felt that both the game design project and the game design presentations were valuable and improved their learning. They were more neutral about the direct benefit of adding peer evaluation to the game design project, although we hypothesize that their engagement in the development of rubrics used to grade the project may have improved their perception of the game design presentations. Most striking is that these benefits were seen in a challenging course that heavily emphasizes writing and has large non-major enrollments.

5. REFERENCES