Educational Game Effectiveness: Answering the Call

Neil M. Redfield
Jay Science and Engineering Academy
United States of America
Neil.Redfield@earthlink.net

Diane Gaither
Carol L. Redfield, Ph.D.
Computer Science Department
St. Mary's University (in San Antonio)
United States of America
DianeGaither@gmail.com
CRedfield@stmarytx.edu

Abstract: An effectiveness study was performed at John Jay Science and Engineering Academy in San Antonio, Texas of a computer game to see if playing the game can positively impact students' knowledge and performance. This paper reports on the process of developing the study of an algebra 1 computer game and the results from that study. *Math Blaster Algebra by Davidson/Knowledge Adventure* was given to students who were in algebra 1. This paper includes how the game was selected, what kind of data was collected, and what analysis was performed. The most significant result we found was a *9% increase of test scores for the group that used the game compared to the control group*. Math Blaster Algebra made the most difference in understanding and manipulating linear equations. Some further initiatives are outlines for studying the effectiveness of computer games and their impact in the classroom.

Introduction

Games are a popular pastime around the world. Kalah from Africa, Mahjongg from China, Reversi or Othello from the United Kingdom, Go from Japan, and Backgammon from Roman times are example of these games. Today, we can play many games by computer either on our machines directly or over the Internet. Games have been used for educational purposes since before we had classrooms. Consider jacks, hopscotch, and concentration where hand-eye coordination, counting, and memory are not only needed but exercised. While these games have been around for centuries, the idea of electronic educational games had not been realized until more recent times.

The use of computer software in education and training dates to the early 1940s, when American researchers developed flight simulators that used computers to generate simulated-onboard instrument data (Wikipedia 2006). The arrival of the personal computer in 1975 changed the purpose and usage of software, allowing specific implications for educational software. By the early 1980s, the availability of personal computers allowed for the creation of companies and nonprofit organizations that specialized in educational software. An immense number of titles were developed and released from the mid-1990's onwards, aimed primarily at the home education of younger children (Gaither, Redfield 2006). The design of educational software programs for home use has been influenced strongly by computer gaming concepts – in other words, they are designed to be fun as well as educational.

Teachers who used educational games in their instruction made some generalizations and subjective observations regarding how well some games have impacted a child's education. However, until 1999, there were no well-known formal studies with any statistical significance. In 1999, a survey was performed on what CD-ROMs were available to teach or practice concepts in K-12 curriculum and adult education. This research found that there was only one software tool that had an effectiveness study performed, Reader Rabbit by The Learning Company now owned by Riverdeep (Gaither, Redfield 2006).

In the summer of 2005, the Games That Teach database was compiled with over 1000 educational computer games. The database is searchable at www.wingz2fly.com/GameSurvey. The games range from all subjects including mathematics to the arts, all ages from toddler to elder, and all methods such as elemental presentation and complex simulation. In the paper Survey of Games That Teach, presented at the AACE/SITE conference March 2006, it was noted there were no effectiveness studies found on any current-market educational games (Gaither, Redfield 2006), although Prensky (2001) mentions that there have been many effectiveness studies performed, the studies must have been done in the research labs and on tutoring systems. We did find references to studies on computer-based training systems that showed the computer tutors are effective in having students learn (Goettl, Halff, Redfield, Shute 1998). This AACE/SITE 2007 paper is intended to answer the challenge made in this previous paper to perform an effectiveness study on an educational game.

We had access to some high schools that might be willing to host a study. We thought that mathematics would be a good subject to select in the study since it is fairly consistent in the content that is taught. Algebra 1 was selected since all students are required to pass algebra and there were a number of software titles to choose from. John Jay Science and Engineering High School in San Antonio offered support and access to algebra 1 students. Neil Redfield is a student there and he met with the students in their algebra 1 classes.

We considered a number of first-year algebra software titles including the following products:

Software Title	**Publisher**
Algebra Animator	Riverdeep
Algebra Concepts	Ventura Educational Systems
Algebra Stars	Sunburst
Algebra World	Math Realm
Math Blaster Algebra	Davidson/Knowledge Adventure
Mighty Math: Astro Algebra	Riverdeep
Mind Power Math: High School	The Learning Company/Riverdeep
Quickstudy Algebra 1	Selectsoft Publishing
Windows Algebra	ProOne

We wanted a software tool that would run on current versions of Windows XP, covered much of the algebra 1 curriculum, was easy to use, was in our budget, and provided a game environment.

Many of the software titles did not run on the Windows XP systems we tried. Many of the games were not available to purchase the 60 copies that we needed. A few of the games were too expensive to consider for our effectiveness study given our limited budget. We had three professors from St. Mary's University in San Antonio's Computer Science and Mathematics Departments evaluate the potential games for usability and algebraic accuracy. We also had a 9th grade student try out the game to see if the games would be considered fun. *Math Blaster Algebra by Davidson/Knowledge Adventure won.*

On the game packaging, the publisher of Math Blaster Algebra says that the game provides tools to succeed and improve skills in a student's first year of algebra (Knowledge Share 1998). The skills include

understanding algebraic expressions and equations,
working with ratio,
proportion and percent,
plotting points on a graph,
factoring polynomials,
applying the order of operations,
exploring inequalities and quadratic equations, and
building and solving equations.

Math Blaster Algebra focuses on using many types of numbers including decimals, integers, both positive and negative, and rational numbers.

The Study

We looked at a number of factors and scores to determine if playing the game made a difference in the leaning process. An algebra 1 test was developed by a certified math teacher and reviewed by the other math professors at St

Mary's University in San Antonio. This test consists of 25 multiple choice questions and was used as a pre-test and post-test. The topics covered in the test were selected by reviewing the Texas Essential Knowledge and Skills and looking at what was included in Math Blaster Algebra. The topics include

comparing expressions,
simplifying expressions (multiple variables, roots, fractions),
solving for a variable,
substituting values into functions,
the domain and range of a function,
translations,
combining like terms,
greatest common factors,
factoring quadratics,
equations of lines,
absolute value,
inequalities,
solving quadratic equations,
graphing linear equations,
solving systems of two linear equations, and
word problems that require an understanding of these concepts.

We had two groups of high school students participating in the study. An invitation was made to 90 ninth grade algebra 1 students in John Jay Science and Engineering Academy. Half of the 42 students who self-selected to participate were given a copy of Math Blaster Algebra to play at their home for 5 weeks. Another group participated in the study by testing and providing grades and scores, but not playing the game, to serve as a control group. A pre-test was administered to all students before the start of the game-playing period. Students in the test group were asked to play the game at least 2 hours a week, however they ended up playing an average of 3.25 hours a week during the study period. Following the five-week game-playing period, we administered a post-test, identical to the pre-test, to all the students. We evaluated these tests to determine the progress that the students made in participating with the game. We also looked at the grades of the students from the major grading periods. All data was analyzed closely to determine if there was any statistically significant difference between the students who played the Math Blaster Algebra game from those who did not play the game.

Results

Both the pre-test and the post-test were the same 25 questions. They were graded by awarding 1 point for correct answers, subtracting ¼ a point for incorrect questions (as the SAT is scored), and no penalty for unanswered questions, making for a maximum score of 25 and minimum score of -6.25. The students pre-test scores ranged from -0.25 to 9.75 for the test group and -2 to 11.25 for the control group. The scores on the post-test ranged from -0.5 to 13.75 for the test group and -1.75 to 14.75 for the control group.

The pre and post-tests measure the entire algebra 1 curriculum for a complete year. The study took place between weeks 10 and 15 of a total of 36 weeks of classes in a year. The concepts the students should have already known and the concepts that were covered in the class room over the study period amount to 10 of the 25 questions of the pre and post-test.

Twenty-six t-tests were performed and included in the analysis on the data gathered from the study. The t-Tests performed were

(A) Pre-Test Scores, Between Groups
(B) Post-Test Scores, Between Groups
(C) Pre-Test to Post-Test Difference, Between Groups
(D) Test Group Pre-Test to Post-Test Difference
(E) Control Group Pre-Test to Post-Test Difference
(F) Pre-Test Correctly Answered Questions, Between Groups
(G) Pre-Test Incorrectly Answered Questions, Between Groups
(H) Post-Test Correctly Answered Questions, Between Groups

(I) Post-Test Incorrectly Answered Questions, Between Groups
(J) Test Group Score Only Using Objectives Covered
(K) Control Group Score Only Using Objectives Covered
(L) Test Group Percent Answered Correctly of Objectives Covered
(M) Control Group Percent Answered Correctly Objectives Covered
(N) Pre-Test to Post-Test Score Difference for Objectives Covered
(O) Pre-Test to Post-Test Percent Answered Correctly Difference for Objectives Covered
(P) Test Group 1st Grading Period to 2nd Grading Period
(Q) Control Group 1st Grading Period to 2nd Grading Period
(R) Test Group 1st Grading Period to 3rd Grading Period Progress Report
(S) Control Group 1st Grading Period to 3rd Grading Period Progress Report
(T) 1st Grading Period to 3rd Grading Period Progress Report Difference Between Groups
(U) 1st Grading Period Averages, Between Groups
(V) 2nd Grading Period Averages, Between Groups
(W) 3rd Grading Period Progress Report Averages, Between Groups
(X) Test Group 1st Grading Period to 2nd Grading Period Benchmark
(Y) Control Group 1st Grading Period to 2nd Grading Period Benchmark
(Z) Grading Period Benchmark to 2nd Grading Difference, Between Groups

The following paragraphs discuss any significant findings were for each of the data tests. Any tests not mentioned showed no significant results.

- According to t-Test (A), the control group, on average, did nearly 12% better than the test group. Based on this difference and the P(T<=t) value (0.018), we can conclusively say the control group's scores for the pre-tests were significantly higher than the test group's scores for the pre-test.

- In the t-Test (B) comparing the post-test scores, the control group did barely 1% better than the test group. Since the P(T<=t) (0.4) is above 0.05, we can assume there was no significant difference between the test group and the control group's post-test scores.

- The t-Test (C) that compares the pre-test to the post-test difference shows us several things. First, the test group, on average, did 3% better on the post-test than the pre-test. Second, the control group did about 6% worse on the post-test than the pre-test. This difference could be caused by the fact that the post-test was given the day after a week-long vacation, so, perhaps, the test group actually would have done just as bad as the control group had they not been playing Math Blaster Algebra over the study period. Assuming this reasoning is the case, playing the game caused *an increase of 9% in the post-test score for the test group over the control group.*

- Paired t-Test (J) shows the test group's scores from the 10 questions that should have already been covered from both tests. The table's values show no significant difference in the comparison of the pre-test and the post-test scores, since the P(T<=t) is .20.

- Paired t-Test (K) showed the control group's scores of only the 10 questions previously mentioned. The P(T<=t) is 0.004, therefore there was a significant decrease of almost 4% from the pre-test to the post-test.

- T-Test (L) is a paired t-Test on the percentage of correctly answered questions on the 10 questions the subjects should have already known at the current time in the algebra 1 class. There was an increase from the pre-test to the post-test. However, since the P(T<=t) (0.19) is greater than 0.05, this difference is not significant.

- T-Test (M) compared percentage correct from the pre-test to the post-test in the control group. There is a visible drop from the pre-test to the post-test, and according to the P(T<=t) (0.01), this drop is significant.

- T-Test (N) compared the pre and post-test score difference between groups when only grading the questions which contained content the students were exposed to. According to this data test, there is a

significant difference, and if it is assumed that the test group's score would be the same as the control group's, then it can also be assumed that Math Blaster Algebra caused an increase of 6% or 1½ points.

- T-Test (O) compared the percentages of correctly-answered questions of the content students were previously exposed to between groups. This shows a significant increase. It can also be assumed that playing Math Blaster Algebra caused an increase of 17% correctness.

Conclusions

Depending on which area of the data was examined, results were very different. Looking at the comparison of the 1st Six Week Averages and the 3rd Six Week Progress Reports, both the test group and the control group, overall, had no significant change. The data is not ideal as it does not take into account all the assignments and projects done over the study period and later material may have been more difficult for the students than material covered earlier.

Regarding the pre-test and post-test scores, the control group had significantly higher scores than the test group on the pre-test, but for the post-test score, there was no significant difference. The combination of the fall in the control group and the rise of the test group shows a 9% difference or increase for the test group.

The analysis of the 10 questions covering concepts that the students should have been exposed to either before and during the study period showed there was no significant difference for the test group from the pre-test to the post-test scores, while the control group had a significant decrease. It was shown that the test group had an increase of 6% in score. When the 10 questions were examined individually, *Math Blaster Algebra made the most difference in understanding concepts of linear functions especially identifying their characteristics and manipulating variables.*

Each grading period consists of 6 weeks of classroom study. The middle of each grading period (each 3 weeks) a progress report is issued to the students to indicate their grade so far. The 1st grading period and 3rd grading period progress report averages for all students participating in the study were collected, as well as scores from two tests during the study period. The report card scores for the 1st grading period were finalized before the project started. The 2nd grading period averages were finalized halfway through the study, and the 3rd grading period progress report was issued when the project ended. The two paired t-tests on this data showed the same results, that there was no significant difference from the end of the first grading period to halfway through the 3rd grading period.

This project was a valuable learning experience for not only the study members, but for us as well. One of the students even mentioned that her pre-school brother now knows what an integer is because he was watching her play. Among other things, we learned that not every effect of these games can be measured. We also learned that not all human subjects are the most responsible, so for anyone planning to do a similar study, remember to set extra time for the returning any forms. We'd also like to suggest having several areas of data that are comparable; you don't know what factors may be affecting them.

It appears that Math Blaster Algebra helps students' understanding of specific algebraic topics. It has been shown that the time on task can make just as much positive difference in student learning than good computer-based tutors. (Hall 1987) *Maybe the real value such educational games provide is a motivation and incentive for spending more time on the task of learning and playing around with whatever topic that is embedded in the game.*

This study was only a preliminary one. There are many changes that could =have been made to this project to improve on it, such as expanding on the study period and expanding the subjects that are participating. There is an extension study that is currently being done at John Jay Science and Engineering High School with many of the same students and some new students. It will be running the entire second semester as opposed to only 5 weeks. We would like to see expanded research into these areas as well as considering

- monitor playing time
- vary playing time
- compare to tutoring by a person
- multiple games
- other subject areas
- multiple schools

Acknowledgements

There were many influential people in this study. First and foremost we would like to thank Ms. Monica Gonzales, the amazingly generous algebra teacher who not only gave us countless hours of class time to be with the students, but also compiled a great amount of data for us to use. We would also like to thank Ms. Peggy Greff, the principal of the magnet school inside of John Jay High School, who game us the go-ahead to use her students. Thanks to all the administration of John Jay Science and Engineering for their support. Thanks to Mr. Mike Scully who oversaw the pieces of the project that were presented in the school science fair (the project is going on to the Junior Science Academy and the Alamo Regional Science and Engineering Fair in the spring of 2007). Finally, we thank the students who participated in the study. Obviously, without them, there would be nothing to report.

References

Gaither, Diane (2006). Survey of Electronic Games that Teach. 2006. Accessed 25 Oct 2006 http://www.wingz2fly.com/GameSurvey/index.html

Gaither, D., Redfield, C. (2006). "Survey of Electronic Games That Teach", Proceedings of the 2006 Society of Information Technology and Teacher Education International Conference, Orlando, Florida: Association of the Advancement of Computing in Education March 2006

Goettl, B., Halff, H., Redfield, C., Shute, V. (Eds.) (1998). *Intelligent Tutoring Systems*, Proceedings of the 4th International Conference on Intelligent Tutoring Systems. Springer-Verlag, Berlin, Germany.

Hall, Douglas (1987). "A Comparative Analysis of Guided vs. Query-Based Intelligent Tutoring Systems (ITS) Using a Class - Entity - Relationship - Attribute (cera) Knowledge Base", University of North Texas (and UMI Proquest), Denton, Texas, 1987

Kafai, Yasmin (2001). The Educational Potential of Electronic Games: . 27 Oct. 2001. University of Chicago. 25 Oct 2006 http://culturalpolicy.uchicago.edu/conf2001/papers/kafai.html

Knowledge Adventure (2006). "Math Blaster" Accessed 28 Jan 2007. www.knowledgeadventure.com

Knowledge Share (1998). "SuperKids Educational Software Review." Accessed 28 Jan 2007. www.superkids.com/aweb/pages/reviews/math/algebra/1/blaster/merge.shtml

Prensky, Marc (2001). Digital Game-Based Learning. McGraw Hill, New York.

Wikipedia (2006). History of Board Games. 25 Oct. 2006. Wikipedia. Accessed 25 Oct 2006 http://en.wikipedia.org/wiki/History_of_board_games