Using Design Experiments As a Means of Guiding Software Development

Steven McGee
NASA Classroom of the Future
Wheeling Jesuit University
316 Washington Avenue
Wheeling, WV, 26003
mcgee@cet.edu

Bruce C. Howard
NASA Classroom of the Future
Wheeling Jesuit University
316 Washington Avenue
Wheeling, WV, 26003
howard@cet.edu

Abstract: The following report illustrates how educational researchers and software developers can work in conjunction with teachers to optimize learning and provide an empirical basis for software revision. Using the Design Experiment Technique, teachers and researchers together determined, through successive classroom implementations, how to optimize the use of new software. Objective assessments made along the way indicated whether the curricular adjustments led to improved learning. The current research involved four implementations of software that teaches science content and inquiry skills. After each implementation, the data were analyzed, and classroom activities were revised. Revisions included the addition of time tables, the jigsaw method for doing background research, daily questions, and activity summaries. These revisions will also be incorporated into future versions of the software.

Introduction

User feedback is a central component of a successful software development project, whether the project is targeted for entertainment, self-help, or other commercial purpose. A primary consideration is ease of use, and developers customarily solicit impressions from users or observe users as they attempt to run the software. Developers of educational software, on the other hand, are interested not only in ease of use, but also in the impact the software will have on learning. Unfortunately, the difficulty of obtaining accurate user feedback on learning is complicated by the diversity of implementations teachers may use to integrate the software into their curricula. The research reported here illustrates how educational researchers and software developers can work in conjunction with teachers to optimize learning and provide an empirical basis for software revision.

The Design Experiment Technique is a recent advancement in educational research (McGee & Howard, 1998). Using this technique, teachers and researchers together determine how to optimize the use of new software through successive classroom implementations. For each implementation, there are objective measures that assess whether the curriculum adjustments have led to improved learning outcomes relative to the previous implementation. This technique was used by researchers at the NASA Classroom of the Future to evaluate the *Astronomy Village®: Investigating the Universe*™ multimedia package. This paper will describe the design experiment process for four classroom implementations and how that process led to refinements.

Astronomy Village: Investigating the Universe

The NASA Classroom of the Future™ Program (COTF) is a NASA-funded research and development center that specializes in the development and testing of educational multimedia for math, science, and technology education. In March 1996, the COTF produced a CD-ROM called *Astronomy Village: Investigating the Universe* for use as a

curriculum supplement in high school science classrooms. It was distributed to over 11,000 teachers, educators, and resource centers, and it won *Technology and Learning* magazine's Science Software of the Year Award for 1996. *Astronomy Village* uses the metaphor of living and working at a mountain-top observatory (the village) as the primary interface from which students investigate contemporary problems in astronomy (see Pompea and Blurton, 1995). Academic activities are designed to promote learning of astronomical concepts and processes related to scientific inquiry. Students join a research team and choose one of ten investigations to complete. In the Stellar Nursery investigation, for example, students investigate how stars are born. For each investigation, students progress through five phases: background research, data collection, data analysis, data interpretation, and presentation of results. For any given phase, there are from three to seven content-related activities to be completed before proceeding to another phase. The primary means of tracking progress through an investigation is the *Research Path Diagram*—a chart that displays icons representing activities within each phase. Each time a student clicks on one of the icons in the Research Path Diagram, a virtual mentor appears and describes activities relevant to that particular investigation. A second means of tracking progress is the electronic LogBook, in which students record their scientific notes and observations.

Method and Data Sources

The design experiment consisted of four one-semester studies. In each study students from one of two area schools visited the COTF facilities to use the *Astronomy Village* software.

Design Experiment Populations

The participating schools were from a rural community with a population of approximately 35,000. The demographics for each study varied (see Table 1). In the first study thirteen students from the ninth-grade class of a girls' academy (college preparatory) participated. In the second study nine students from the eighth-grade class of the same academy participated. In the third study twelve students from the tenth and eleventh grade of a large public high school participated. In the fourth study nine students from the eigth-grade class of the same girls' academy participated. Students from the third study were from an at-risk population. In each case students attended class daily for approximately four weeks at the COTF facility in lieu of their science class. Sessions using *Astronomy Village* were co-taught by the students' classroom teacher and the first author.

It should be noted that the students in Studies 2 and 4 were younger than students in Studies 1 and 3, and in Study 3 the population consisted of at-risk youth in a public school setting. Research predicts that the ninth-grade students from the college preparatory academy should have outperformed the other groups of students (those in lower grades and at-risk) (Blank and Gruebel, 1995).

Data Sources

There were three sources of data for this investigation. The first source was a compilation of documents in the students' electronic notebooks. In the first two studies students used the electronic notebook embedded within *Astronomy Village*. This notebook was called the LogBook. In the third and fourth studies, students used an electronic notebook called the Collaboratory Notebook as their LogBook. The Collaboratory Notebook was developed as part of the Learning Through Collaborative Visualization Project (CoVis) Project at Northwestern

	Study 1	**Study 2**	**Study 3**	**Study 4**
Time Frame	Apr-May, 1996	Oct-Nov, 1996	Mar-Apr, 1997	Oct-Nov, 1998
Num. of Students	13	9	12	9
Gender	All Female	All Female	F = 4; M = 8	All Female
Grade Level	9th	8th	10th-11th	8th
Type of Students	college prep	college prep	at-risk	college prep

Table 1: Demographics of Design Experiment Populations

University (Edelson and O'Neill, 1994). For each activity in *Astronomy Village*, students were expected to produce activity summaries in their LogBooks. These summaries consisted of a brief description of the activity, a statement of the importance of the activity in the context of the main research question, and any new questions that arose from the activity. The second source of data was videotapes of student interactions while using the software. The third source of data was field notes and classroom observations by the teachers and first author.

Dependent Measures

There were two aspects of the activity summaries that were considered for each design experiment. First, how many of the activities in *Astronomy Village* did students complete? Within each study the primary dependent measure for this question was an indicator of activity completion called the *activity completion rate.* This was an objective measure of the extent to which students were able to complete the assigned activities. Activity completion rate was defined as the number of activity summaries students completed, divided by the number of activities the mentor suggested. Second, were students able to relate the individual activities to the basic question for each investigation? In each study the primary dependent measure for this question was a measure of how closely the statement of importance within the activity summaries was linked to the overall investigation.

Design Experiment Results

In brief, the design experiment began by using the instructional procedures prescribed in the *Astronomy Village* package, and examining the effects as the program was implemented. Each succeeding study refined the procedures in order to achieve more effective outcomes. The next section describes each study and the impact of changes to the activity on the activity completion rate. Along with a description of each successive study, there is a discussion of how the instructional procedures were modified for the next implementation.

Activity Completion Rate

Study 1

In the first study the overall goal for the students was to complete all of the activities as suggested by the *Astronomy Village* virtual mentor. Our purpose was to implement the program as closely as possible to the intentions of the software designers. Students completed activities related to background research, data collection, data analysis, data interpretation, and presentation of results. As intended by the developers, students began by watching a videotape introduction to the software. Next, they worked within their project teams to complete a tutorial on using the software. After completing the tutorial, students proceeded to the virtual Conference Center in *Astronomy Village*, where the virtual mentors described the different investigations. The students selected different mentors and listened to a description of the research that each mentor was conducting. The students chose an investigation that interested them and joined an investigation team. Next, they accessed the Research Path Diagram, which introduced them to the resources on the *Astronomy Village* CD-ROM that related specifically to their investigation. By following the Research Path Diagram, students should have been able to complete the steps necessary to learn the appropriate concepts and conduct the level of problem solving needed for their investigation.

Completion Rate. Summary worksheets made activity completion easier, although completion rates were still below teacher expectations. The average activity completion rate for this study was 42%. This value indicated that over the four-week period, students completed less than half of the activities the mentor suggested. Further analysis of the activity completion rate within each phase of research (background research - 55%, data collection - 75%, data analysis - 35%, data interpretation - 20%, and presentation - 27%) revealed that the activity completion rate was much lower during later phases of research. This analysis of the activity completion indicated that the Research Path Diagram and virtual mentors were not sufficient to guide students through the complex problem solving required by the activities in the investigation. If students were not able to complete the activities in the investigation, it was believed they would not achieve the desired learning outcomes.

Study 2

In considering the disappointing performance of the students in Study 1, the teachers for Study 2 modified the implementation slightly from that intended by the designers of the *Astronomy Village.* The modification took three forms: *target dates* for phase completion, *more time* for activity completion, and *more contextualized training*.

Modifications. First, students were given target dates for each of the phases of research so they would have sufficient time to complete later phases of research. Second, teachers eliminated activities that did not seem to benefit the students. That is, the tutorial was eliminated, and students selected a investigation from a list of abstracts instead. These two changes resulted in approximately 17% more time for completion of activities (3.5 instructional days out of 20). Third, training was now given on an as-needed basis.

Completion Rate. The average activity completion rate for this study was 85%. This value indicated that over the four-week period, students completed twice as many activities as the students in the first study, even though these students were a year younger. The activity completion rate within each phase of research was as follows: background research - 78%, data collection - 100%, data analysis - 83%, data interpretation - 62%, and presentation - 100%. The students in this study had more opportunity to achieve the desired learning outcomes than students in the previous study. Thus, the use of an objective measure such as the activity completion rate made it possible to see what factors were preventing students from completing the activities.

Study 3 & 4

Although the modifications in Study 2 improved the completion rate, the teachers noticed that students were having difficulty seeing the connection between the individual activities and the investigation as whole. In Study 3, the researchers and teachers initiated several major modifications to help students see the connection. These modifications took three forms: all students working on the same investigation, the addition of a motivating question phase, and a shortened investigation structure.

*Modifications***.** To help students complete their investigations more quickly, the research phases were revised and truncated. The new phases were the motivating-question phase, the background research phase, the background review phase, the data analysis phase, and the reflection phase. In the motivating-question phase the teacher posed the main investigation question, and the students individually typed responses in their LogBooks. Next, the teacher showed the students the data they would be analyzing and asked them to record observations. These two activities were meant to activate students' prior knowledge and connect it to the activities of the investigation. This approach was similar to Minstrell's benchmark lessons (Bruer, 1993). In the background research phase the teacher selected the three most relevant articles from the investigation, and each student read one of the three and developed an activity summary. In the background review phase students used their activity summaries to answer questions as a team, prompting themselves to integrate across the individual readings. Since each student was an expert on only one of the articles, students would be forced to discuss the readings with one another in order to answer the questions. In the data analysis phase students completed analysis worksheets as a team. And finally, in the reflection phase students responded to integrative, teacher-posed questions in their LogBooks. Using these redesigned research phases, it was possible to fit three investigations, rather than one, into the four-week period.

Completion Rate. The average activity completion rate for Study 3 was 74%. This rate was comparable to the activity completion rate from Study 2, even though the third group consisted of students who were at risk. The activity completion rate for each phase was as follows: motivating question - 82%, background research - 85%, background review - 62%, reflection - 73%. The activities for this study involved more explicit prompts for students to reflect on their background knowledge and relate that knowledge to the activities in the research investigation. These more explicit prompts increased the likelihood that students would assimilate the new information from *Astronomy Village* into existing knowledge structures.

Study 4 used the same structure as in Study 3, with one primary exception. The students only completed one investigation during an eight-day period. Also, researchers once again worked with eighth-grade students.

LogBook Quality

In addition to activity completion rates, researchers wanted to know how well students were connecting their activities over the four-week period to the larger context of the basic science question. (A basic science question addressed a fundamental issue in astronomy such as, Is there other life in the universe? In contrast, a research investigation question addressed a specific issue in astronomy such as, How do you determine what areas of the universe to examine for evidence of planets orbiting a distant star?) To evaluate these connections, the researchers examined students' LogBook responses to a particular question in the activity summary worksheets that queried the importance of the activity. A rating scale was developed to judge the responses. Responses that related the activity to a basic science question were rated more highly than those that merely related the activity to the research investigation question, while those that discussed only the activity itself (zero connection) were rated lowest. Statements of importance were extracted from the LogBook activity summaries for each article that students read. Raters were blind as to which study the statement of importance came from. Raters made judgments by comparing the statements of importance to the article content, the investigation research question, and the basic science question.

Researchers investigated the impact that the various implementations had on the quality of the LogBook entries. In order to simplify the analysis, the studies were clustered into 2 phases, based on the similarities of the implementations. Phase 1 encompassed Studies 1 & 2, in which students worked on one investigation *at their own pace* over four weeks. Phase 2 encompassed Studies 3 & 4, in which students worked on several investigations-- as directed by the teacher-- and used a truncated set of activities for each investigation. In addition, Phase 2 studies were more explicit in providing students with assistance in making connections between activities.

Results indicated that students from Phase 2 studies were able to connect almost one-third of the activity summaries to the basic science question for each investigation (see Figure 1). In contrast, students from Phase 1 studies were able to connect roughly one-third of their activity summaries *only* to the research investigation question. The two phases were comparable in the percent of activities that contained no connection between the articles and the research investigation or basic science question. A Chi-square analysis indicated that the higher percentage of students in Phase 2 who were capable of connecting the activity summaries to the larger science context was statistically significant ($\chi^2 = 12.19$, $p < .01$).

These results indicated that the addition of a motivating-question phase and a reflection phase helped the students to draw connections between the various activities within the extended project. By making explicit references to the basic science question and prompting students to develop hypotheses about the basic science question, the teachers in Phase 2 helped students connect individual activities to the larger scientific context.

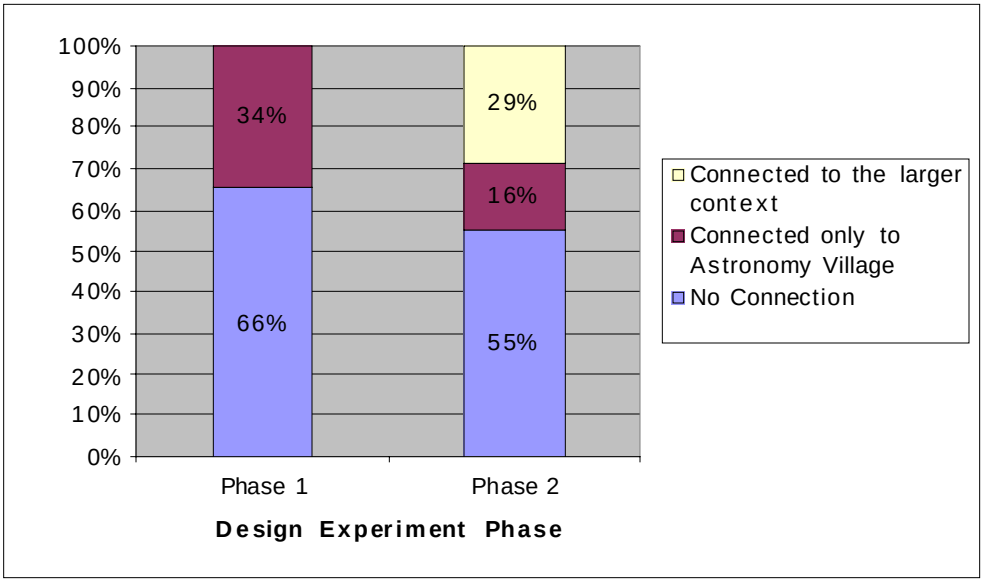

Figure 1. Percent of Activity Summaries that Relate to the Larger Context of the Basic Science Question

Conclusions

The specific modifications in this design experiment were unique to the *Astronomy Village* software and to the particular classroom setting used at COTF. However, the general issues that were revealed are germane to many software programs and classroom contexts. If teachers understand the issues they will face while implementing technology and are given guidance on how to address these issues, there will be a greater chance that educational reform will be sustained beyond the initial pilot projects (McGee & Howard, 1998).

The results of the design experiment helped the developers at the COTF understand the kinds of supports that would be needed to optimize learning. These supports are now being incorporated into a new version of *Astronomy Village* that will focus on the solar system. In this new version the investigations will be shortened to one week to help students better manage their time. With more time, students will have four opportunities to conduct investigations in a four-week period. To help students connect individual activities to the investigation question, an exploration phase will also be added, to take place prior to the background research phase. The exploration phase will serve the purpose of the motivating question phase in Study 3. In addition, the LogBook will provide structured prompts (similar to the activity summaries) to help students reflect on the learning process.

Reform will only take place when teachers are able to explore the implications of a new curriculum within the context of their own classrooms. That an educational multimedia program was used successfully in a one school district does not necessarily mean that it will be effective in another. Design experiments can be especially powerful when results lead to teacher-generated design principles. Creating partnerships between teachers and researchers will enhance reform by giving teachers tools for evaluating their own implementations and providing researchers a means of comparing implementations across different contexts.

References

Blank, R. K., & Gruebel, D. (1995). *State Indicators of Science and Mathematics Education 1995*. Washington, DC: Council of Chief State School Officers.

Bruer, J. T. (1993). *Schools for thought: A science of learning in the classroom*. Cambridge, MA: The MIT Press.

Edelson, D. C., & O'Neill, D.K. (1994). The CoVis Collaboratory Notebook: Supporting collaborative scientific inquiry. In A. Best (Ed.), *Proceedings of the 1994 National Educational Computing Conference (pp.146-152)*, . Eugene, OR: International Society for Technology in Education in cooperation with the National Education Computing Association.

McGee, S. & Howard, B. (1998). Evaluating educational multimedia in the context of use. *Journal of Universal Computer Science*, 4(3), 273-291. [Online] Available: http://www.cet.edu/research/papers.html [1999, March 25].

Pompea, S. M., & Blurton, C. (1995, Jan-Feb). A walk through the Astronomy Village. *Mercury*, 32-33.

Technology and Learning (1996). Vol 17, #3.

Acknowledgements

This research was supported in part by grants from the National Aeronautics and Space Administration (NCCW-0012) and by the National Science Foundation (ESI-9617857). We would like to thank the teachers and students who participated in the design experiment. We would like to thank Brigitte Gegg for her assistance in conducting the studies and processing the data. We would like to thank John Hornyak and Steven Croft for answering our endless questions about astronomy content. We would like to thank all of our Wheeling Jesuit University student interns who helped in analyzing the data. We would like to thank Dorothy Frew for their helpful comments on earlier drafts of the paper.