Evolution of Academic Tasks in a Design Experiment of Scientific Inquiry

Steven McGee, Bruce C. Howard, & Namsoo Hong
NASA Classroom of the Future,
Wheeling Jesuit University, Wheeling, WV
Paper Presented at the Annual Conference of the
American Educational Research Association, San Diego, CA, 1998

Abstract

Reported here are preliminary results from classroom implementations of a recently developed program titled *Astronomy Village: Investigating the Universe*, produced by the NASA Classroom of the Future program (COTF). Over a period of two school years, academic tasks for engaging students in scientific inquiry were designed and redesigned based on the results of previous implementations.

Theoretical Framework

Over the last century, science educators have tried repeatedly to reform science education. From decade to decade the goal has been the same: to help students engage in scientific inquiry and cease being passive receptors of information. In every era of reform, there have been pilot projects that have successfully engaged students in scientific inquiry. Examples include project-based curricula from the Progressive Era (Collings, 1923; Tanner, 1997) and the new science curricula of the Sputnik Era (Klopfer & Champagne, 1990). In such projects, students learned more and developed healthier attitudes toward science than did students in traditional science classrooms (Shymansky, Kyle, & Alport, 1983). However, such projects have not become the norm in spite of their promise.

According to Cuban (1993), reform efforts often fail to sustain at the classroom level because they largely ignore the inevitable fact that innovative curricula provided to teachers must be adapted. In their efforts to accommodate the often conflicting demands of many constituencies such as school boards, administrators, parents, and students, teachers cannot use all instructional techniques and materials exactly as intended. Real-world constraints necessarily factor into how teachers implement educational innovations in their classrooms. Teachers are forced to become instructional designers, making decisions about what features of the innovation to adapt and what features to preserve (Ball & Cohen, 1996). However, if teachers make the wrong adaptations, if they overlook or distort the critical elements of the innovation, not only will students fail to engage in scientific inquiry, but the entire reform movement itself will be judged by the failed implementations. Therefore, it is imperative for designers to articulate the critical elements of a particular innovation.

More recently, reformers have begun to investigate the nature of how teachers implement reform efforts in the classroom. Projects such as the Third IEA Mathematics and Science Study (TIMSS) and QUASAR have articulated the distinction between the intentions of the curriculum designers and the implementation of the curriculum by the teachers. The TIMSS project characterizes the distinction as the *Intended Curriculum* and the *Implemented Curriculum*. The QUASAR project goes further, distinguishing between the curriculum as *Set-Up* by the teacher and the curriculum as *Implemented* by the students (Stein, Grover, & Henningsen, 1996). Curriculum activities as set-up by the teacher are linked to the curriculum activities as implemented by the students through the concept of “academic tasks” (Doyle, 1983). Doyle defines an academic task by the *goals* that students are expected to accomplish, the *resources* that are available to students while they are pursuing the task goals, and the *operations* that students can use in manipulating the resources to accomplish the task goals. The goals as assigned by the teacher and the resources made available to the students influence the types of operations that student can engage in. This paper will discuss a design experiment that investigates the evolution of academic tasks and the resulting impact on student performance.

The Multimedia Evaluation Target: 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, COTF published a CD-ROM called *Astronomy Village: Investigating the Universe* for use as a curriculum supplement in high school science classrooms. It has been distributed to over 11,000 teachers, educators, and resource centers, and 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 both 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 shows each phase of the investigation and icons to represent activities within each phase (see *Figure 1*). Each time a student clicks on one of the icons in the Research Path Diagram, a virtual mentor appears and describes activities that are relevant to the investigation. Students also have access to an electronic LogBook for recording their scientific notes and observations.

Figure 1. Stellar Nursery research path diagram

Method and Data Sources

The design experiment consisted of three studies, each covering one semester. In each study students from one of two area schools used the COTF facilities to conduct astronomy investigations using the 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. 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.

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

Table 1: Demographics of Design Experiment Populations

It should be noted that the students in Study 2 were younger than students in Study 1, and in Study 3 the population shifted from college preparatory students in a private school to at-risk youth in a public school setting. Research predicts that the ninth-grade students from the college preparatory academy should outperform the other groups of students (Blank and Gruebel, 1995).

Data Sources

There were three sources of data for this investigation that enabled academic task research to support the design experiment. 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 study, students used an electronic notebook called the *Collaboratory Notebook,* developed as part of the Learning Through Collaborative Visualization Project (CoVis) Project at Northwestern University (Edelson and O'Neill, 1994). 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.

For each academic task in *Astronomy Village*, students were expected to produce written responses in their electronic notebooks. Examples of notebook entries included activity summaries, answers to "press conference" questions, answers to teacher-posed questions, and reflections on "thought experiments." Activity 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. (For the pilot research reported here, only statements of importance were examined, and only those that were responses to completing an article.) Press conference questions were posed to students by members of a simulated press corps. Students responded with answers from their investigation. Thought experiments prompted students to integrate path activities. Other collected virtual data included images, snippets of articles students had read, or student reflections.

Dependent Measures

Prior research on academic tasks has shown that it is essential that students be able to complete academic tasks in a cognitively demanding way in order for them to learn from the tasks (Hiebert and Wearne, 1993; Stein and Lane in press). Within each study the primary dependent measure was an indicator of task completion called the *task completion rate.* This was an objective measure of the extent to which students were able to complete the task goals for the assigned academic tasks. Task completion rate was defined as the number of activity summaries students completed, divided by the number of academic tasks the mentor suggested.

The task completion rate was mediated by the level of cognitive task demand associated with each academic task. In utilizing academic tasks, teachers must consider the cognitive demands of the task in the context of learning objectives. In order for learning to take place, there needs to be an appropriate amount of cognitive demand. If an academic task is too demanding, students will encounter a cognitive overload and will have difficulty completing the task (Sweller and Chandler, 1994). If an academic task is not demanding enough, students won't learn from the task. Through an examination of the task completion rate, it became possible to identify academic tasks that might be too demanding for the students. In those cases where students were systematically not completing a task, the teachers provided instructional supports that ostensibly lessened the cognitive overload associated with those tasks. Student notebooks, teacher field notes, and video tapes provided the data to make judgments about task demands.

The second dependent measure was an evaluation of the extent to which the students' academic task operations mirrored the nature of scientific inquiry. Researchers wanted to know whether a) students adopted the research question for a given pathway in *Astronomy Village*; b) students were able to build connections between individual academic tasks and the overall research question of the pathway; and c) students generated ideas or explanations for these connections. For the pilot analyses reported here, only b) was examined— the degree to which students connected individual activities to the larger context.

Design Experiment Results

In brief, the design experiment began with using the prescribed instructional procedures intended in the *Astronomy Village* curriculum, and examined the effects as the curriculum was implemented. Each succeeding study refined the procedures in order to achieve more effective outcomes. The next section describes each study, the academic tasks used, and their impact on the task completion rate. With each successive study, there is a discussion of how instructional procedures were modified for the next implementation.

Activity Completion Rate and Task Demand

Study 1

In the first study (20 days in April-May, 1996), the overall goal for the students was to complete all of the academic tasks as suggested by the virtual mentor. Our purpose in this study was to implement the curriculum as closely as possible to the curriculum as intended by 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 were available to describe the different investigations. The students selected different mentors and listened to a description of the research that each mentor was conducting at the Village. 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. Each recommended activity was defined as one academic task for purposes of this evaluation.

Within the first week of software use, the interplay between task completion and task demand became apparent. Students were expected to record the results of each academic task in their LogBooks. However, the software did not provide any templates to help the students create LogBook entries. The students were either copying and pasting entire articles into their LogBooks, taking detailed notes, or not writing anything in their LogBooks. The teachers felt that none of these strategies were going to be effective for the students to be able to synthesize all of the investigation activities. Therefore, the teachers created within the LogBook an activity summary worksheet that students used to summarize each activity. For each summary, the students were asked to include a brief description of the activity, a statement indicating how the activity fit into their investigation, and any new questions that arose during the activity. The activity summary worksheet provided an instructional support to lessen the cognitive overload associated with the investigation.

Activity summary worksheets made task completion easier, although completion rates were still below teacher expectations. The average task completion rate for this study was 42%. This value indicates that over the four-week period, students completed less than half of the activities that the mentor suggested. Further analysis of the task completion rate within each phase of research (i.e., background research - 55%, data collection - 75%, data analysis - 35%, data interpretation - 20%, and presentation - 27%) revealed that the task completion rate was much lower during later phases of research (see *Figure 2*). This analysis of the task completion rate and task demand indicates that the Research Path Diagram and virtual mentors were not sufficient to guide students through complex problem solving required by the tasks in the path. If students were not able to complete the tasks in the path, it was believed they would not achieve the desired learning outcomes.

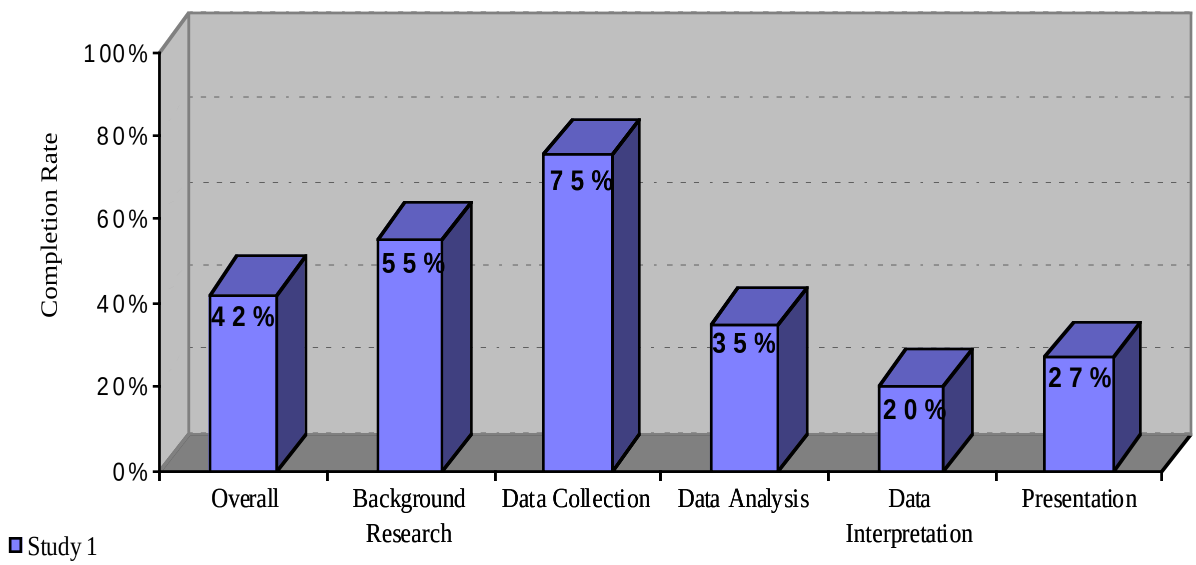

Figure 2. Task Completion Rate by Research Investigation Phase for Study 1

In examining additional data from the four weeks of instruction, there were several noteworthy observations. First, students spent most of their time in the background research phase of *Astronomy Village*, which left little time for the remaining activities. This imbalance was most likely due to the high task demand of background research activities, which hindered students' attention to time management issues. That is, students were observed to be engrossed in these early activities, to the exclusion of process-related discussions or questions that would have indicated that they were tracking individual activities in the context of the larger investigation. Second, students averaged 3.5 class periods to complete the tutorials and select their investigation. Student comments and questions indicated that what was learned during the advance tutorials was not remembered very well, if at all, when the time came to apply such knowledge. It was concluded that such tutorials, if needed at all, should be completed on an as-needed basis, thereby lessening cognitive overload during the earlier phases, where demand was highest. In the next study, the teacher and first author used these results to redesign task demands to improve the task completion rate.

Study 2

In considering the disappointing performance of the students in Study 1, the teachers for Study 2 (20 days in October-November, 1996) modified the curriculum slightly from that intended by the designers of the *Astronomy Village*. The goal of the modification was to relieve cognitive overload so that task completion could be improved. The modification took three forms: target dates for phase completion, more time for task completion, and more contextualized training.

First, students were given target dates for each of the phases of research so that they would have sufficient time to complete later phases of research. This instructional support was intended to lessen cognitive overload during the phase of background research. Second, the teachers eliminated activities that did not seem to benefit the students. That is, students selected a pathway from a list of abstracts prior to beginning the study, and the tutorial was eliminated. These two changes resulted in approximately 17% more time for completion of academic tasks (3.5 instructional days out of 20). Third, training was given in the context of use, which was intended to lessen cognitive overload incurred by needing to remember various software procedures that would be used in later phases of research. Instead, the training was done at the beginning of each phase of research by the teacher demonstrating software features that would be needed.

The average task completion rate for this study was 85% (see Figure 3). This value indicates 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 task 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 task completion rate made it possible to see what factors were preventing students from completing the academic tasks.

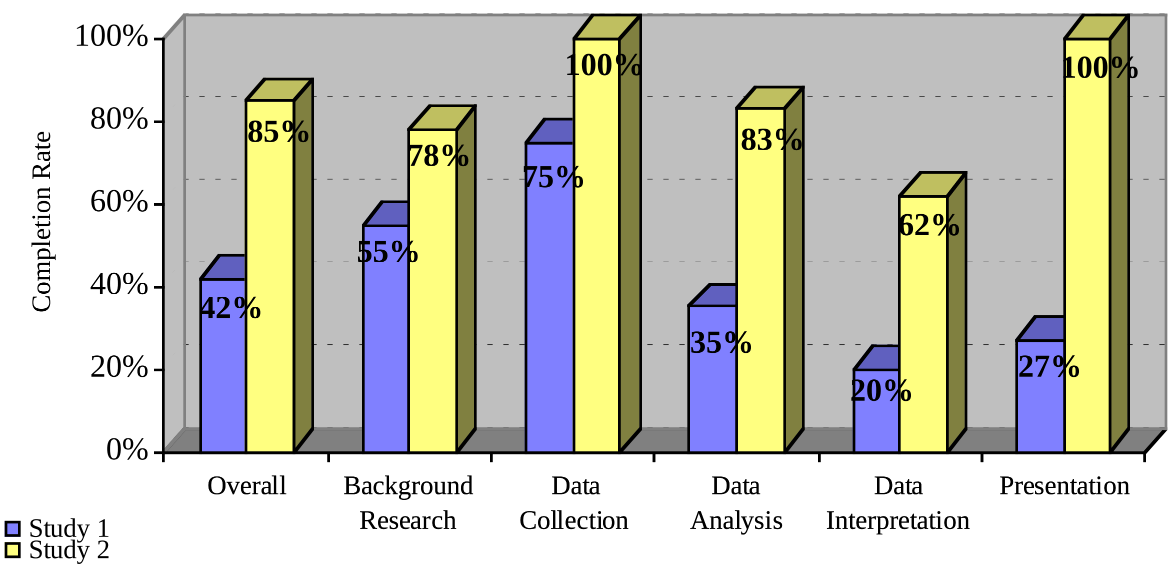

Figure 3. Comparison of task completion rate across Studies 1 and 2

At the end of the second study, researchers reviewed the videotapes, LogBooks, and field notes from both of the studies. Two important observations were made. First, within the activity summary, it was noted that students were capable of developing brief descriptions of the tasks. However, they were not developing good statements of how the activity related to the overall research path. Second, there was no evidence that students were activating prior knowledge about the research question in order to build connections to these new experiences. The emphasis of the third study was on helping students build connections between prior knowledge, academic tasks and the research question.

Study 3

In Study 3 (19 days in March-April, 1997), the researchers and teachers initiated several major modifications. The purpose of all modifications was to provide more structure to support the activation of prior knowledge and provide opportunities for students to complete multiple investigations in four weeks. Rather than have students select the paths to work on, the teacher and first author selected the pathways that students would work on, and all student teams investigated the same pathways.

In order to accomplish the goal of completing investigations more quickly, research phases were revised and truncated into five alternative phases: the motivating question phase, background research, background review, data analysis, and reflection. In the motivating question phase, the teacher posed the main investigation question and the students individually typed responses in their electronic notebooks. Next, the teacher showed the students the data that 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 pathways, and students each read one of the three articles and developed an activity summary. In the background review phase, students used their activity summaries to answer questions as a team that would prompt students to integrate across the readings that were done individually. Since each student was an expert on only one of the articles, students would be forced to discuss the readings with each other in order to answer the question. 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 notebooks. Using these redesigned pathways, it was possible to fit three investigations paths into the four-week period.

The average task completion rate for Study 3 was 74%. This rate was comparable to the task completion rate from Study 2, even though the third group consisted of students who were at-risk. The task completion rate for each phase was as follows, motivating question - 82%, background research - 85%, background review - 62%, reflection - 73% (see Figure 4). The academic tasks for this study involved more explicit prompts for students to reflect on their background knowledge and relate that knowledge to the tasks in the research path. These expanded prompts increased the likelihood that students would assimilate the new information from *Astronomy Village* into existing knowledge structures.

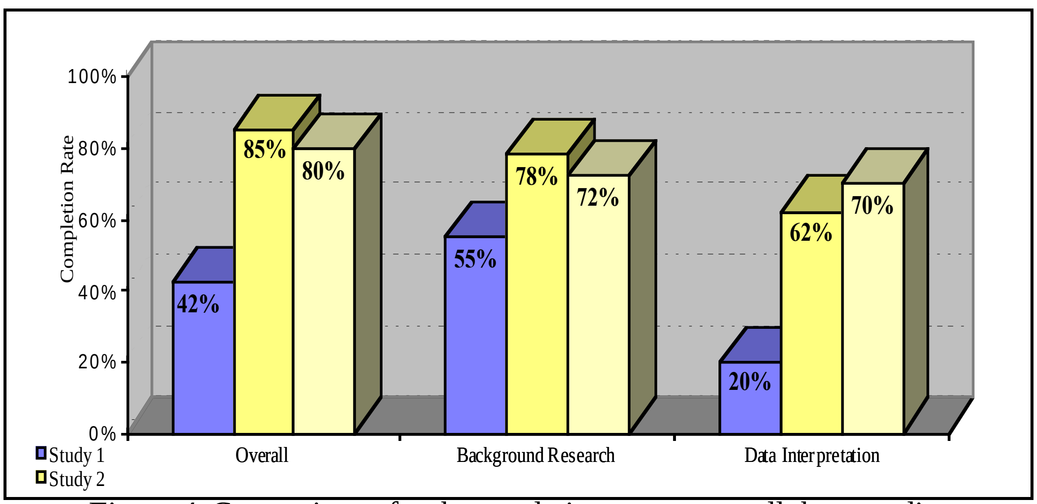

Figure 4. Comparison of task completion rate across all three studies

Number of Activities Completed

Perhaps the most remarkable finding of this design experiment was the comparison of the number of activities completed across the three studies. Students in Study 3, although using a comparable amount of instructional time, completed nearly twice the number of activities. For the pilot analyses reported here, only the article reading activities were examined. Overall, it was found that students in Study 1 finished 16 articles, students in Study 2 finished 15 articles, and students in Study 3 finished 40 articles (see Figure 5)

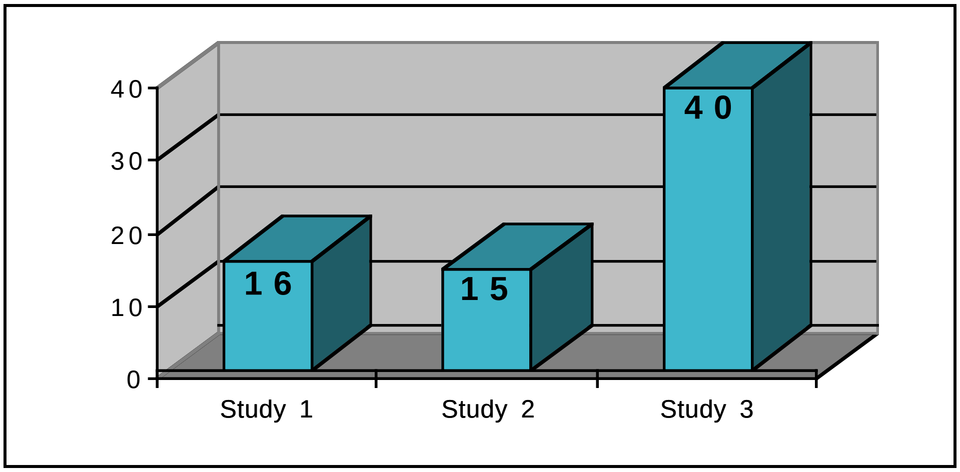

Figure 5. Number of Articles Read

Activity Significance: Connecting to the Big Picture

In addition to task completion rates and task demand evaluations, 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' 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.

For the pilot analyses reported here, only summaries for reading articles were examined. "Connection scores" were based on the following rubric:

4 = contains an idea from the article AND an idea related to the basic science question
3 = contains an idea from the article AND an idea related to the path research question
2= contains only an idea related to the basic science question or path research question
1 = contains only an idea from the article
0 = no scientific concepts or ideas expressed

Statements of importance were extracted from the activity summary 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 path research question, and the basic science question. Inter-rater reliability was a=.85, F=2.45 (p=.12), R=.73 (p<.0001).

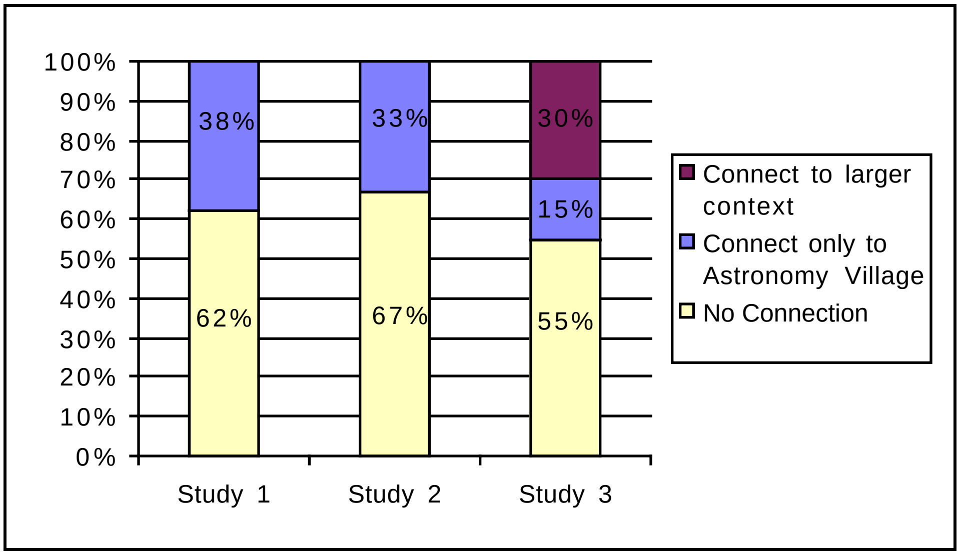

Figure 6. Percent of Activity Summaries that Relate to the Larger Context of the Basic Science Question

Figure 6 indicates that students from Study 3 were able to connect almost one-third of the activity summaries to the basic science question for each pathway. In contrast, students from Study 1 and 2 were able to connect roughly one-third of their activity summaries only to the research path question. The three studies were comparable in the percent of activities that contained no connection between the articles and the research path or basic science question. A Chi-square analysis indicated that the higher percentage of students in Study 3 who were capable of connecting the activity summaries to the larger science context was statistically significant ($c^2 = 12.71$, $p < .05$). These results support the hypothesis 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 Study 3 helped students connect individual activities to their larger scientific context.

Summary of Results

Students were able to complete more activities when provided more structure. Students who were provided with a streamlined curriculum (Study 3), which lessened the need to know procedural knowledge associated with the computer-based curriculum, were able to complete over twice the amount of article research. The evidence indicated that students in the streamlined curriculum made more connections to the larger context of the basic science question. This finding was particularly important since students in Study 3 were from at-risk backgrounds and had reported motivational problems. Students in the other two studies were from a private, elite, college preparatory girls' school.

The main findings of this study were that students need structure in processing individual activities in the context of a multi-week curriculum centered around researching a singular scientific question. By providing this structure, along with deadlines for task completion, students were able to finish more and understand more. In the third study, by scaffolding conceptual development through building links to prior knowledge, the teacher was able to maintain task completion rate while altering the level of task demand and cover more investigations with a more difficult population Considering these findings, it was recommended that future implementations of this curriculum include the following: scaffolding and fading (such as activity summaries), activating prior knowledge (such as asking motivating questions), and coaching procedural knowledge directly (such as contextualized training for using software features).

Conclusions

If teachers are to use curriculum materials as a means to reform their practices, it is important for developers to communicate the manner in which the materials could be adapted to different classroom contexts. The specific adaptations in this design experiment were necessarily unique to the COTF context. However, the general issues that were revealed through the design experiment are germane to many classroom contexts. If teachers understand the issues they will face when implementing technology and are given guidance on how to address these issues, there will be a greater chance that the reform project will be sustained beyond the initial pilot projects.

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 neighboring school district does not mean that it will be effective in a new teacher's classroom. 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.

Future Directions

When designing instruction, teachers need to balance the level of cognitive challenge of an academic task with students' abilities to complete the task. The most cognitively demanding tasks are those that are most difficult for students to complete. With the task completion rate indicator, teachers can identify those parts of an academic task that students are having the most difficulty with. In the case of *Astronomy Village*, the task completion rate indicator served to alert the teacher and researcher to the fact that students were not engaging in efficient planning of their solution, preventing them from completing academic tasks that came later in the project cycle. Once alerted to the problem, the teacher made adjustments in Study 2 that increased the task completion rate without negatively affecting the cognitive demand of the task. Future research at COTF will focus on developing efficient indicators for monitoring cognitive task demand.

Just as the results of each individual study led to new hypotheses and promoted instructional redesign, the results of the design experiment led to working hypotheses for the next experiment. COTF will use the results of this study in further curriculum evaluations to design models of appropriate task demands and to develop assessment instruments to examine the curriculum as intended and implemented

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.

Collings, E. (1923). *An experiment with a project curriculum*. New York: The Macmillan Company.

Cuban, L. (1993, October). The lure of curricular reform and its pitiful history. Phi Delta Kappan, 182-185.

Doyle, W. (1983). Academic work. *Review of Educational Research*, 53(2), 159-199.

Doyle, W., & Carter, K. (1984). Academic tasks in the classrooms. *Curriculum Inquiry*, *14*(3), 129-149.

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.

Hiebert, J., and Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal*, 30(2), 393-425.

Klopfer & Champagne 1990

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

Shymansky, J. A., Kyle, W. C., & Alport, J. M. (1983). The effects of new science curricula on student performance. *Journal of Research in Science Teaching*, 20(5), 387-404.

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*(2), 455-488.

Stein, M. K., & Lane, S. (in press). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation.*

Tanner, L. N. (1997). *Dewey's Laboratory School: Lessons for Today*. New York: Teachers College Press.

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