ExoQuest® Beta Test Report

Jennifer Kirby
Steven McGee, Ph.D.
Center for Educational Technologies®
Wheeling Jesuit University

Dimiter Dimitrov, Ph.D.
Kent State University

1. Background

In the ExoQuest® software program students travel on virtual journeys to destinations in the solar system and beyond. Their trips are based on past, present, and future NASA astrobiology missions. At each destination students conduct investigations that include hands-on and simulated experiments. Each investigation poses problems that focus on different areas of research, providing an interdisciplinary approach to science and the scientific method.

There are two types of activities in ExoQuest, background modules and mission modules. Background modules cover general topics that help prepare students for the mission modules. Background modules introduce students to concepts and techniques that are explored in more detail in the missions.

The ExoQuest beta test was conducted during spring 2001. The treatment group consisted of 15 teachers with 748 students, while the control group consisted of 7 teachers with 260 students. Control group teachers and students were from the same schools as the treatment group. Treatment group teachers were asked to recruit another teacher in their school with class distributions close to their own classes. The treatment group administered the 45-question test before and after using ExoQuest, while the control administered the same test during the same time frame without using the software between administrations. Both groups had slightly more females (T = 57.3%, C = 51.3%) than males (T = 42.7%, C = 48.7%) involved in the testing. The treatment group students were from grades 6-12, while the control group had students in grades 7-8 and 10-12. Table 1 below shows the percentage breakdown of grade for both groups.

Table 1

	6	**7**	**8**	**9**	**10**	**11**	**12**
Treatment	2.5%	19.6%	31.3%	33.8%	5.1%	4.9%	2.7%
Control		8.8%	87.3%		2.7%	.8%	.4%

2. Pre- and Posttest Results

The Astronomy Village® Search for Life test was adapted for use with the ExoQuest project. The main enhancement was the addition of nine publicly released items from National Assessment of Educational Progress and the Third International Mathematics and Science Study selected by the project manager. The Search for Life test was designed to test students' understanding of the core requirements of life. In particular, it focused on water and an energy source. The underlying

approach was to test students' ability to transfer what they learned in the instructional environment to problems that they would not have encountered in that environment. The modules should help students improve their performance on the content portion of the test. The Search for Life test also measured how well students can evaluate evidence of planetary conditions to determine whether life might exist there. The test focused on image analysis of dry riverbeds, analysis of fossil evidence, infrared image analysis, and temperature-pressure relationships.

Students were given a 45-question pre- and posttest, of which 21 were basic content questions and 24 were problem-solving questions. Table 2 shows the pretest and posttest means for control and treatment groups. A repeated measure ANOVA shows that there was a statistically significant interaction between pretest/posttest and control/treatment groups for the content portion of the test. There were no statistical interactions present for the problem-solving portion.

Table 2

Pretest and Posttest Means		
	Pre	**Post**
Treatment		
Content	59.75%	62.20%
Problem Solving	53.70%	55.50%
Control		
Content	56.25%	56.09%
Problem Solving	46.85%	46.12%

Content – $F(1,1649) = 4.52$, $p < .05$

As reported in other research, there can be some serious drawbacks to using raw pretest and posttest scores to measure a change in learning outcomes (Dimitrov et. al., 2001). First, raw scores do not adequately reflect the underlying students' ability on a given test. Second, a repeated measure ANOVA does not separate treatment and trend effects. Treatment effects are attributed to some treatment administered between the pretest and posttest. Trend effects can be attributed to factors such as natural biological maturation and cognitive development over the period of time between pretest and posttest.

We used an item response theory technique call the linear logistic model for change (LLMC) to analyze trend effects due to ExoQuest. The LLMC models student response patterns to provide a more accurate representation of underlying abilities then raw scores. In addition, LLMC is able to separate treatment and trend effects. In this context the term ability connotes a latent trait that underlies the student's performance on a test (e.g., Hambleton, Swaminatan, & Rogers, 1991). It does not imply general ability. The units of the ability scale, called "logits," typically range from –4 to +4. They represent natural logarithms of odds for success on the test item. If a student succeeds on 75 percent and fails on 25 percent of the test items, the odds ratio for the test is 3/1 = 3. Thus the ability score of this student is the natural logarithm of 3, which is 1.10 (i.e., about one unit above zero on the logit scale). Other basic concepts are treatment effect and trend effect that measure ability changes due to treatment and natural trends, respectively. With the LLMC

design, the treatment effect is set to zero for the control group, but both experimental and control groups may have a nonzero trend effect because of factors such as natural biological maturation and cognitive development over the period of time between pretest and posttest measurements.

The LLMC results in Table 4 show that there was no statistically significant trend effect on content understanding *(*0.011, p<.05) or problem solving (0.046, p>.05). There was also no statistically significant treatment effect for the group on either of the two scales, content understanding (0.132, p>.05) and problem solving (0.041, p>.05). Consistent with the ANOVA model, there were no trend effects. However, when the trend effect was separated from the treatment effect and there was a better model for the underlying students' abilities, we no longer see a treatment effect.

Table 4

Pretest to Posttest Treatment and Trend Effects for ExoQuest

	Content Understanding		Problem Solving	
Treatment	Treatment effect	Trend effect	Treatment effect	Trend effect
ExoQuest	0.132 (0.072)	0.011 (0.064)	0.041 (0.046)	0.046 (0.039)

Note. The standard errors of the effect estimates are given in parentheses. All treatment and trend effects are on the same ration scale (in logits). The ratio of any two effects indicates how many times one effect is greater (or smaller) than the other effect.

Overall, students tend to pick up some of the basic concepts of exobiology. However, the results indicate that the software did not support students in synthesizing these concepts to be able to transfer them to new contexts.

A content analysis of ExoQuest shows that little emphasis was placed on helping students to connect individual activities to the larger context. Previous research indicates that connecting individual activities to the larger context is a difficult task for students, but it is also necessary for development of complex understanding.

3. Activity Evaluation – Multiple Choice Questions

In addition to learning outcomes, CET researchers sought student and teacher opinions about the program. Both teachers and students were asked to evaluate the program using an online form (see Appendix A) after they completed each section of ExoQuest. The form consisted of four multiple choice and two open-ended questions relating to the students' perception of each activity. There were a total of 1,987 valid evaluations entered. Each evaluation was linked to the particular mission and activity that the students were working on. The results were summarized

by both missions and activity type. Averages were taken to determine the value of each question for each category (see appendix B for complete breakdown).

On the opened-ended section students were asked to answer two questions: 1) What do you like best about this activity? 2) What changes would you recommend need to be made to this activity? After reading through all the responses, the most popular suggestions are listed below. Appendices C and D show a list of all responses.

- **Mission Modules**
  - How I liked it (1= not at all, 4=a lot)
    - Overall average = 2.55
    - Liked Remote Sensing the most (3.01)
    - Liked Extrasolar Planets the least (1.91)
  - What I learned (1= didn't learn anything, 4= a lot)
    - Overall average = 2.69
    - Perceived to have learned the most on Extreme Environments (3.10)
    - Perceived to have learned the least on Extrasolar Planets (2.09)
  - Ease of instructions (1= very difficult, 4= very easy)
    - Overall average = 3.03
    - Easiest instructions were Extreme Environments (3.41)
    - Hardest instructions were Drake-ulator (2.23)
  - Ease of activity (1=very hard, 5= very easy)
    - Overall average = 3.25
    - Easiest activity was Stuff of Life (3.60)
    - Hardest activity was Drake-ulator (2.47)

- **Activities**
  - How I liked it (1= not at all, 4=a lot)
    - Overall average = 2.64
    - Liked Research Guide the most (2.93)
    - Liked Briefing Video the least (2.23)
  - What I learned (1= didn't learn anything, 4= a lot)
    - Overall average = 2.79
    - Perceived to have learned the most on Research Guide (3.17)
    - Perceived to have learned the least on Briefing Video (2.42)
  - Ease of instructions (1= very difficult, 4= very easy)
    - Overall average = 3.16
    - Easiest instructions were Introduction (3.47)
    - Hardest instructions were Sci-Logs (2.76)
  - Ease of activity (1=very hard, 5= very easy)
    - Overall average = 3.40
    - Easiest activity was Briefing Video (3.76)
    - Hardest activity was Sci-Logs (3.11)

4. Activity Evaluation – Open-Ended Questions

- **Likes**
  - All the new interesting information learned.
  - Activities were pretty easy to follow.
  - Video helped explain what the mission was about.
  - Loved interactive hands-on activities.
  - Coming up with our own ideas and brainstorming.
  - I liked how we got to talk in a group about things..
  - I liked guessing.

- **Dislikes**
  - Need to have clearer instructions and questions. Didn't know when completed with an activity or what to do next.
  - There is too much reading for the children to do.
  - More visuals and interactive activities.
  - Videos move too fast.
  - I did not learn anything from doing this, mostly because we did find out if we had put down the correct answers or not. Therefore, we did not know if what we put down was right so we couldn't tell if we had learned anything.
  - Have it actually hands on; while it was nice to pick things, that was all we got to do.
  - Have less videos and more interaction.

Table 5 shows the correlations for the activity evaluation responses. With such a large dataset it is somewhat easy to obtain significance, so we were more interested in the strength of the correlation rather than significance level. As the results indicated, whether or not the students liked the activity is highly correlated to how much they felt they learned. Also, the ease of the activity was highly correlated with their perception of the clarity of instructions.

Table 5

	Like	Perceived Learning	Instructions
Perceived Learning	**.6871 (.000)**		
Instructions	.2941 (.000)	.2550 (.000)	
Easy/Difficult	.0527 (.019)	.0109 (.627)	**.3946 (.000)**

5. Conclusion

The ExoQuest program provides a rich context for exploring exobiology. As seen by the activity evaluations, both teachers and students responded favorably to the program. However, future revisions should provide support for the larger context. The above results show that while students were able to increase their content knowledge of exobiology, they were not as successful with their problem-solving skills. Students were not able to transfer their knowledge from the learning environment to a new context.

6. Appendix A

ExoQuest® Student Activity Evaluation

1) How I liked it--
- I did not like it at all.
- I liked it very little.
- I liked it some.
- I liked it a lot.

2) What I learned--
- I did not learn anything at all.
- I learned very little.
- I learned some.
- I learned a lot.

3) About the instructions--
- The instructions were very difficult to follow.
- The instructions were somewhat difficult to follow.
- The instructions were somewhat easy to follow.
- The instructions were very easy to follow.

4) How easy/difficult--
- This activity was very hard.
- This activity was a little hard.
- This activity was just right.
- This activity was a little easy.
- This activity was very easy.

5) Individual or group--
- I did all of it by myself.
- I did most of it by myself.
- I did most of it with others.
- I did all of it with others.

6) If you worked in a group--
- We had a lot of difficulty working together.
- We had some difficulty working together.
- We worked somewhat well together.
- We worked very well together.

7) What did you like best about this activity?

8) What changes would you recommend need to be made to this activity?
(Also report any problems you may have had.)

7. Appendix B

ExoQuest® Activity Evaluation by Module					
	N*	**How I liked it**	**What I learned**	**About the instructions**	**How easy/ difficult**
To Whom It May Concern	348	2.72	2.89	3.02	3.15
Distance and Time	299	2.83	2.80	3.10	3.55
Requirements for Life	276	2.76	2.89	3.15	3.35
Drake-ulator	160	2.16	2.39	2.23	2.47
Remote Sensing	131	3.01	3.06	2.89	3.20
Stuff of Life	194	2.62	2.73	3.22	3.60
Water in the Solar System	102	2.48	2.70	3.12	3.21
Extreme Environments	164	2.94	3.10	3.41	3.43
Ocean on Europa	92	2.11	2.35	3.10	3.13
Viking Exobiology	141	2.68	2.83	3.22	3.32
Extrasolar Planets	11	1.91	2.09	2.55	3.00
ET Communications	69	2.36	2.45	3.32	3.55
Overall Average	1987	2.55	2.69	303	3.25

* N= number of activity evaluations per module

ExoQuest Activity Evaluation by Activity					
	N*	**How I liked it**	**What I learned**	**About the instructions**	**How easy/ difficult**
Briefing Video	122	2.23	2.42	3.33	3.76
Introduction (in Sci-Corder)	195	2.82	3.08	3.47	3.56
Assignment Video (in Sci-Corder)	230	2.85	2.96	3.40	3.37
Research Guide (in Sci-Corder)	164	2.93	3.17	3.38	3.52
Simulation	687	2.75	2.82	2.94	3.12
Hands On Activities	152	2.39	2.45	2.86	3.33
Sci-Log	437	2.48	2.61	2.76	3.11
Overall Average	1987	2.64	2.79	3.16	3.40

* N= number of activity evaluations per activity