Evaluation of Seeds of Science/Roots of Reading Project: Shoreline Science and Terrarium Investigations

CSE Technical Report 676

Jia Wang and Eva L. Baker
CRESST/University of California, Los Angeles

March 2006

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Center for the Study of Evaluation (CSE)
Graduate School of Education & Information Studies
University of California, Los Angeles
GSE&IS Building, Box 951522
Los Angeles, CA 90095-1522
(310) 206-1532

The work reported herein was supported in part by the University of California, Berkeley (via the National Science Foundation), Award SA4576-10056

The findings and opinions expressed do not reflect the positions or policies of the National Center for Education Research, the University of California, Berkeley, or the National Science Foundation.

TABLE OF CONTENTS

EVALUATION OF SEEDS OF SCIENCE/ROOTS OF READING PROJECT:

SHORELINE SCIENCE AND *TERRARIUM INVESTIGATIONS*

Jia Wang and Joan Herman

CRESST/University of California, Los Angeles

Abstract

This project was initiated in order to evaluate two literacy and science integrated instruction units, *Shoreline Science* and *Terrarium Investigations*, designed by the Lawrence Hall of Science *Seeds of Science/Roots of Reading* Project (*Seeds/Roots*). We examined how the integrated units affect student interest, motivation, and learning, as well as evaluating the units' quality, usability, and utility using both quantitative and qualitative approaches. Through analyses of student performance and teacher interviews, we found *Shoreline Science* and *Terrarium Investigations* beneficial to both students and teachers. Teachers were highly motivated to use the materials, and Shoreline Science and *Terrarium Investigations* students learned significantly more than the control group students in all science and literacy measures on which the differences were expected.

The Lawrence Hall of Science (LHS) *Seeds of Science/Roots of Reading* Project (Seeds/Roots) is funded by the National Science Foundation (NSF) to design integrated science and literacy instructional units. Developed collaboratively by experts in science education and literacy development, the project draws on Great Explorations in Math and Science's (GEMS) highly regarded inquiry-based science content to create a coherent curriculum addressing the development of primary grade students' science and literacy skills. The project thus faces the complex challenge of coordinating appropriate developmental continua and devising materials, guidance, and effective sequences that systematically support students' development in both subject areas. An innovative leap for curriculum development, Seeds/Roots implementation will also require new learning and perspectives for teachers. The evaluation of the integrated instruction units must address these challenges, in addition to more common questions of curriculum implementation and impact. Shoreline Science and Terrarium Investigations were the first two units developed and field-tested under the Seeds/Roots project. They were piloted with

elementary students in Grades 2 though 4 in the school year 2004-2005. The pilot included 25 teachers using Shoreline Science and 19 teachers implementing Terrarium Investigations in their classrooms.

The National Center for Research on Evaluation, Standards, and Student Testing (CRESST) was asked to evaluate how the integrated units affect student interest, motivation, and learning, as well as evaluate the quality, usability, and utility of these first two units in the series. The CRESST evaluation plan used both quantitative and qualitative approaches. Quantitatively, we evaluated the effects of Shoreline Science on student achievement scores in science and literacy separately, using available data. Qualitatively, we conducted a one-hour phone interview with each of the seven Shoreline Science teachers and six Terrarium Investigations teachers who volunteered their additional feedback and opinions about their classroom implementation experiences.

The rest of the deliverable is organized into the following four major sections:

- Evaluation Questions
- Quantitative Evaluation
- Qualitative Evaluation
- Summary

Evaluation Questions

The CRESST evaluation plan will address the following research questions:

1. How do the materials "work?" For example, to what extent and how are the units implemented? How engaged and motivated are students? What are teachers' reactions to the quality, usability, and utility of the units?
2. What problems and/or misunderstandings do teachers and/or students encounter in implementing the materials?
3. What are the effects of using the materials on students' learning of science and reading of informational science texts?
4. For whom, for what purposes, and/or in what contexts do these materials appear most effective?
5. What factors contribute to and/or detract from successful implementation?
6. How can the units be improved? To facilitate implementation? To enhance students' learning in science and reading of informational texts?

The qualitative evaluation will address all the above questions via the teacher interviews. The quantitative evaluation will specifically address question 3 using the models within the hierarchical linear model (HLM) framework. The reasons to use the HLM and the description of HLM will be addressed later in the report.

Quantitative Evaluation

The LHS team recruited teachers through the GEMS network, an International Reading Association listserv, an announcement posted on a web site, and recruitment fliers at the National Science Teachers Association and the International Reading Association. Interested teachers were directed to the LHS web site and asked to apply online. Besides asking for teachers' contact information and a personal statement as parts of the application package, teachers were also asked to provide the following information:

- school community type
- number of years teaching
- number of years teaching at grade level
- grade level
- number of students in the school
- number of students in classroom
- ethnic/racial breakdown of school
- percentage of ELLs in school and class
- hours devoted per week to literacy and science
- names of curricula in use in both subjects
- principal's signature granting permission for participation
- whether or not they would allow video taping in the classroom
- which units they would like to teach

The LHS team received 174 applications in the first round of the recruitment. After careful reviewing of the application packages, they accepted 62 teachers for the studies on Shoreline Science (35 teachers) and Terrarium Investigations (31 teachers; since some teachers wanted to do more than one, the numbers add up to more than 62 teachers.) These teachers were then randomly assigned to either the experimental group using the LHS materials or the control groups. The control group teachers for Shoreline Science would move to be in the treatment group for the Terrarium Investigations study after finishing the Shoreline Science study.

Later, the LHS teamed did a second round of recruitment in order to have two additional internal control groups, the Literacy-only group and the Comparison group, to the Terrarium Investigations study. The Literacy-only group used the nine student readers and a literacy-only curriculum; the Comparison group took the pre- and posttests, but was not supplied with any curriculum materials; and the GEMs group used the Terrarium Habitats GEMS unit. Through the 2nd recruitment, LHS added three teachers to the GEMS group, 12 to the Literacy-only control group, and 10 to the Comparison group. Efforts were made to get a similar pool of teachers for the new group, but the treatment group was somewhat disadvantaged. On average, the teachers in the control groups have more teaching experience, fewer urban and rural students, and more suburban students.

Student Data

At the beginning and end of each unit, a pretest and a posttest in both literacy and science were administered to the students. In Shoreline Science, the pretest and posttest in literacy and science were the same tests. In Terrarium Investigations, the pretests were the shorter versions of the posttests. The items in the pretests were included as part of the posttests on literacy and science, except that the item orderings were different.

Separate data analyses were conducted for the science data and for the literacy data per our study design and to maximize the sample size. We received various student and teacher/school data on Shoreline Science information and Terrarium Investigations information in the period of April to June 2005. The data we received for science and literacy for both Units were similar in terms of the types of information that were available. These are the variables that were available:

1. Student gender (male or female)
2. Student grade level (Grade 2, 3, or 4)

3. Student ethnicity (White, Hispanic, or other)
4. Student home language
5. Student test scores by item
6. Student/teacher group membership (treatment or control)
7. School enrollment size
8. Percentage of students who have limited English proficiency
9. Percentage of students in free or reduced lunch program in a school
10. Percentage of White students in a school

In the Shoreline Science Unit, for the pre- and posttests, all students took an identical science test and an identical literacy test before and after using Shoreline Science. In the science test, there were 19 multiple-choice (MC) items and 11 short-answer (SA) items. In order to investigate whether there were any differences in communicating student understanding in a particular question format versus the other format and whether students behaved differently with MC and SA items, we created three different test scores (overall scores, MC scores, and SA scores) using the available item scores.[1] The outcome variables would be students' gain scores on MC items, SA items, and all items, which were calculated by subtracting students' pretest scores from their posttest scores. The use of gain scores has been widely discussed. One group of researchers claimed that gain scores are not reliable (e.g. Cronbach & Furby, 1970; Lord, 1963) while another group found gain scores to be a good candidate for dependent variables because of the likely reduction of variability related to individual differences (Nicewander & Price, 1983; May & Hittner, 2003). Williams and Zimmerman (1996) argued that the reliability of a test depends on the test construction procedure and the nature of the instrument. We decided to use gain scores accepting the reasoning of Nicewander and Price (1983) and May and Hittner (2003), as well as knowing that students' pretest and posttest were based on the same test.

[1] Using the original science data we received from the LHS, we calculated the reliability coefficients for both pretest and posttest. For the pretest, we based on the calculation on the 1,095 students who answered all 30 questions. The reliability coefficients were 0.7074 for the 30-item test, 0.6223 for the 19-item MC portion, and 0.5941 for the 11-item SA portion. For the posttest, we had a total of 697 students who answered all questions. The corresponding reliability coefficients were 0.7660, 0.6229, and 0.6815, respectively.

The Shoreline Science literacy test consisted of 95 comprehension items and 34 vocabulary items.[2] The comprehension items could be further grouped for each of the four texts as: (a) narrative, (b) target science, (c) science comprehension, and (d) content area comparison.[3] The vocabulary items could be divided into three sub-areas: (a) Shoreline Science only, (b) Terrarium Habitat only, and (c) both Shoreline Science and Terrarium Habitat.[4] Based on the item scores, we created two pretest measures—pre-comprehension scores and pre-vocabulary scores—as control variables in the analysis. For the outcome variables, we used the gain scores in these seven areas, four on comprehension and three on vocabulary items. We analyzed these seven measures separately because different hypotheses were made for each of them. The specific hypotheses are described in detail later.

For the Terrarium Investigations unit, there are nine items in the *Terrarium Investigations* science pretest—4 MC items and 5 SA items.[5] In the science posttest, there were 8 MC items and 9 SA items.[6] We also created three different scores for the MC, SA, and all items. The outcome variables would be students' posttest scores with their pretest scores used as control variables.

The Terrarium Investigations literacy pretest includes 43 comprehension items and 10 vocabulary items. The comprehension items could be further grouped into Danny (18 items) and Life in the Soil (25 items). The vocabulary items are made up of 3 picture vocabulary and 7 association vocabulary items. The posttest had a total of 108 items – 64 comprehension items (18 on Danny, 25 on Life in the Soil, and 21 on the Beach) and 34 vocabulary items (10 picture and 24 association items).

---

[2] Based on 592 students who answered all 129 questions in the literacy pretest, we found the reliability coefficient to be 0.9806 for the 95 comprehension items and 0.9029 for the 34 vocabulary items. Their corresponding coefficients were 0.9854 and 0.9118 for the posttest, based on the analysis of 521 students.

[3] The reliability coefficients were in the range of 0.9310 to 0.9524 for the four groups of comprehension items in the pretest, and the coefficients improved in the posttest for all four groups of items.

[4] The reliability coefficients were in the range of 0.7025 to 0.8370 for the three groups of vocabulary items in the pretest, and these coefficients changed to be in the range of 0.7145 to 0.7799 in the posttest.

[5] For the Terrarium Investigations science pretest, the reliability coefficients were 0.401, 0.578, and 0.583 for the 4 MC items, 5 SA items, and the combined 9 items, respectively.

[6] The c reliability coefficients for the Terrarium Investigations-science posttest items were 0.621 for MC, 0.714 for SA, and 0.714 for all items.

Analysis Methodology

The hierarchical nature of the data—students nested within teachers within schools—made it necessary to choose a methodology that could, in its estimations, take into account this nesting of the data. The widely-used ordinary-least-square regression procedure was not appropriate because it ignores the hierarchical nature of the data, which leads to underestimation of the standard errors and therefore interpretation errors (Burstein, 1980; Raudenbush & Bryk, 2002). Our approach was to analyze the data within the general framework of HLM, which allows the explicit handling of multiple levels of data efficiently and appropriately (Raudenbush & Bryk, 2002). A brief methodological discussion is given below to facilitate the interpretation of the model employed for the estimations. Interested readers should refer to Kreft and de Leeuw (1998), to Raudenbush and Bryk (2002), and to the selected references cited in these books for a more extensive explanation of the conceptual and methodological details of hierarchical modeling.

In a multilevel model, the random variability in the variables observed consists of variability found between the basic smallest unit of analysis, conventionally known as level 1, and variability found between the higher level grouping, level 2. Emphasis is placed on defining and exploring variations at each level and how such variances are related to explanatory variables. In hierarchical models, the possibility of explicitly modeling cross-level interactions or variances associated with the errors at each level of the hierarchy allows for more interesting questions to be asked of the data.

In the present evaluation, the nested nature of the data can be viewed in a hierarchical framework with student observations (level 1) nested in higher level (level 2) groups (i.e., classrooms, teachers, schools). A multi-level model allows the impact of student variables to differ according to the school context (technically, variation in slopes). We chose the hierarchical generalized linear model with two sub-models, namely the level-1 and level-2 models. Specifically, the level-1 model represents the relations among the student level variables and the level-2 model represents the influence of school level factors on the outcome variable through student-level variables (Raudenbush & Bryk, 2002). This is due to the following two factors: (a) we did not have any classroom and teacher variables, and (b) students in each school were taught in the same classroom by the same teacher. In other words, it is technically sound to have the grouping variable at the school level since students in the same school either received the treatment or did not. Since we have

school-level variables available, we decided to investigate the effect of treatment at the school level.

The following has the estimation model we used in the final analysis:

Level 1 equation

$$Y_{ij} = \beta_{0j} + \beta_{1j} \text{PRE-MULTIPLE CHOICE} + \beta_{2j} \text{PRE-SHORT ANSWER} + \beta_{3j} \text{FEMALE} + \beta_{4j} \text{HISPANIC} + \beta_{5j} \text{OTHER} + \beta_{6j} \text{GRADE 3} + \Upsilon_{ij}$$

Level 2 model specifications

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \text{Treatment} + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

In these equations Y_{ij} is the outcome variable for student *i* in school *j*, γ_{01} is the estimated coefficient for the treatment, and u_{0j} is the unique effect of school *j* on the level-1 intercept. The values of γ_{10} to γ_{60} are the estimated coefficients for student level variables, and we assume the effect is consistent across teachers/schools and did not differ by teacher/school.

Please note that the two pretest variables used here are only applicable when we analyze the Shoreline Science science data. For the Shoreline Science literacy, we used literacy pretest scores. And for the Terrarium Investigations unit, we used unit-specific and content-specific pretest variables in the analysis and describe them specifically, when proper, in the later sections.

Quantitative Results for Shoreline Science Science

Twenty-five teachers implemented the newly-designed Shoreline Science materials in their classrooms, and another 10 teachers participated in the study as control group teachers who used the previously-published GEMS materials.

Twenty-one out of the 25 Shoreline Science treatment teachers submitted their data and all 10 control group teachers submitted their data.

Descriptive Results

After excluding students and teachers (and schools) with missing values on the list of variables described above, we had a final sample of 351 students.[7] Table 1 presents the student distribution information as to whether they were in the treatment or control group. Of the 351 students, 248 were in the treatment group and the remaining 103 students were in the control group. The 248 students in the treatment group were taught by 16 teachers at 16 schools, and the 103 control students were taught by 8 teachers at 8 schools.

As shown in Table 1, we had a similar distribution of male and female students in the overall data and by group. The overall data consisted of 60% White students, 28% Hispanic students, and 12% students of other ethnicities. White students were slightly over-represented, and students of other ethnicities were slightly under-represented in the control group. In the control group, 70% of the students were White, 26% were Hispanic, and only 4% were of other ethnicities.

As to grade level, about two thirds of the students were in Grade 3, and the other third were in Grade 2. The distribution of grade level was somewhat different for the control group where there were 46% Grade 2 students and 54% Grade 3 students. The treatment group distribution was similar to the overall data distribution.

The average school enrollment was 449 students; the mean percent of students receiving free and reduced-fee lunch was 49% and the mean percent of White students in each school was 56%. The 16 schools in the treatment group had a slightly higher enrollment size (467 vs. 414), a higher percentage of students (12%) receiving free and reduced-fee lunch, and 3% less White students in their schools than the eight schools in the control group.

[7] In both science and literacy data, we lost students because (a) we could not match students by their ID across the pretest data, posttest data, and background data, (b) about 100 students fewer finished the posttest, (c) some students finished the posttest but not the pretest, (d) four treatment teachers did not submit their data, and (e) some students had missing information on gender and ethnicity.

Table 1

Student Distribution by Student Level Variables for Shoreline Science Science Data

Variable	Value Label	Treatment Group (*N*=248)	Control Group (*N*=103)	Overall (*N*=351)
Female				
	Male	120	51	171
	Female	128	52	180
Ethnicity				
	All Other	37	4	41
	Hispanic	72	27	99
	White, Not Hispanic	139	72	211
Grade Level				
	Grade 2	81	47	128
	Grade 3	167	56	223

Students' mean test scores by their group membership and by item format are summarized in Table 2. Table 2 indicates that students in both groups improved in the posttest, about 3 points (3 items) for MC scores, and about 4 points (4 items) for SA scores. We also found that treatment group students had higher gain scores than control group students by about one point. In other words, when we looked at the gain score differences, treatment group students responded correctly to one more item than control group students did.

Table 2

Student Mean Scores on Shoreline Science Science Pre- and Posttest Tests

		Mean Test Scores		
Item Category	Variable	Treatment Group (*N*=248)	Control Group (*N*=103)	Overall (*N*=351)
Multiple Choice				
	Pretest	12.13	12.63	12.28
	Posttest	16.29	15.80	16.15
	Gain	3.28	2.35	3.01
Short Answer				
	Pretest	2.33	2.41	2.35
	Posttest	5.61	4.76	5.36
	Gain	4.17	3.17	3.87
Combined				
	Pretest	14.46	15.04	14.63
	Posttest	21.91	20.55	21.51
	Gain	7.45	5.51	6.88

HLM Results

Besides examining the data descriptively, we also ran the basic HLM analyses to be certain there were a large number of between-school variations to justify the use of the HLM framework, as compared to the ordinary-least-square regression procedure. The results confirmed that there were significant amounts of variation between schools.[8] With this confirmation, we proceeded to the HLM analysis. We considered including school characteristic variables, such as enrollment, percent of students in free/reduced fee lunch program, and percentage of non-White students together with treatment group status in our school-level estimation model in the initial analyses. After finding all of the school characteristic variables to be non-significant statistically, we excluded them from the final analyses and only included the group membership variable as it was found to be a significant school-level

[8] We ran a basic HLM without any predictors to estimate how the variation found in the outcome variables could be partitioned between students and between schools. We found that 24% of the variation was between schools and 76% was between students for the combined gain scores. The between schools was 17% for the MC gain scores and 18% for the SA gain scores.

predictor. Table 3 has the final results from the HLM analyses for students' SA gain scores, MC gain scores, and MC and SA combined gain scores.

Table 3

HLM Results on Shoreline Science Science Gain Scores (*N*=351)

	Multiple Choice		Short Answer		Combined	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*
School Level						
School average	1.72*	0.84	1.33*	0.48	12.93*	1.16
Treatment	0.67	0.36	0.93*	0.25	1.59*	0.58
Student Level						
Pre-multiple choice	-0.75*	0.05	0.15*	0.03	-0.60*	0.07
Pre-short answer	0.28*	0.08	-0.56*	0.06	-0.29*	0.12
Female	0.39*	0.19	0.35*	0.18	0.73*	0.34
Hispanic	-0.50	0.31	0.05	0.19	-0.47	0.42
Other	0.12	0.31	0.06	0.26	0.20	0.46
Grade 3	0.32	0.55	0.40	0.27	0.86	0.80

*p<.05.

Results on Shoreline Science Science MC Gain Scores.[9] According to the results presented in Table 3, we found that students who used *Shoreline Science* scored higher than control group students, but the difference was not statistically significant. The results also indicated that there were no differences for students of different ethnic groups, and Grade 2 and Grade 3 students scored similarly in their gain scores. Female students, however, improved significantly more than male students did.

For students' pretest scores, their MC pretest scores were negatively related to students' MC posttest scores, and their SA pretest scores were positively related to students' MC posttest scores. We speculated that the negative association between students' pretest MC scores and MC gain scores was an indication that students with lower prior knowledge benefited more from using the Shoreline Science unit

[9] These four student variables explained 49% of the 83% (100% – 17%) of variation we found in the outcome variable at the student level. The indicator variable on whether the teacher (or school) was in the treatment group explained 45% of the 17% of the variation we found at the teacher/school level.

than students who started with more prior knowledge. We also speculated that Shoreline Science materials would "even out" students' initial knowledge differences. We tested our speculation by calculating the mean gain scores for students who were below the mean based on the pretest MC scores and at or above the mean on the pretest MC scores. We found that students who had below mean pretest MC scores gained over five points in the gain score, while students who were at or above the mean pretest MC scores gained two points. This explained the negative association.

Results on Shoreline Science Science SA Gain Scores.[10] The results reported in Table 3 indicated that students who used Shoreline Science improved their SA gain scores by a statistically significant amount as compared to students in the control group. The treatment effect of Shoreline Science improved students' gain scores by one score point (one more item scored as correct) out of the 11-point maximum score. This was the net Shoreline Science effect after controlling for all the student variables in the models and their previous test scores in the pretest. The net effect could be translated into 0.57 effect size. In other words, being a Shoreline Science student improved the SA gain score by 0.57 standard deviations, when compared to non-Shoreline Science students with similar backgrounds and learning contexts.

For students' pretest scores, their MC pretest scores were positively related to students' SA posttest scores, and their SA pretest scores were negatively related to students' SA posttest scores. The same reason could be used to explain the negative association, as we found for the MC gain scores. Students who had below mean pretest SA scores gained more than students who scored at or above the mean in the SA pretest, 3.5 points versus 2.5 points.

There were no statistical differences between students of different ethnic groups, and there were no differences between Grade 2 and Grade 3 students in their gain scores. Female students, however, improved significantly more than male students did in their gain scores.

Results on Shoreline Science Science Total Gain Scores. The results indicated that students who used Shoreline Science improved their total gain scores

---

[10] These four student variables explained 17% of the 82% (100% - 18%) of variation we found in the outcome variable at the student level. The indicator variable on whether the teacher (or school) was in the treatment group explained 53% of the 18% of the variation we found at the teacher/school level.

by a statistically significant amount as compared to students in the control group. The treatment effect of Shoreline Science students improved their test scores by 1.6 score points (one and a half more items scored as correct) out of a 30-point maximum score. This was the net Shoreline Science effect after controlling for all the student variables in the models and their previous test scores in the pretest. The net effect could be translated into about a 0.50 effect size; in other words, being Shoreline Science students would improve students' total gain score by 0.50 standard deviations when compared to non-Shoreline Science students.

Students' total gain scores were negatively associated with their pretest scores for the MC items and SA items. Similarly, we explained the negative associations as the result of students with lower prior science knowledge gaining relatively more from the unit than students with a higher prior knowledge. Students who were below mean in the combined pretest gained 2.5 more points in MC and 0.4 more points in SA than students who scored at or above the mean in the pretest, combining both MC and SA items. There were no statistical differences for students of different ethnic groups, or for students in Grade 2 and Grade 3. Female students, however, improved significantly more than male students did.

Quantitative Results for Shoreline Science Literacy

The literacy test consisted of two main components, comprehension and vocabulary. For the comprehension portion, students were given four texts: (a) "Pack Your Bags" for narrative, (b) "The Beach" for target science, (c) "Life in the Soil" for science comparison, and (d) "Jobs Around Us" for content area comparison. The hypotheses were:

- Treatment group students were not expected to out-perform control students on **narrative** items (25 items).
- Treatment group students were expected to out-perform control students on **target science** items (21 items).
- Treatment group students were not expected to out-perform control students on **science comparison** items, but it would be a great finding (25 items).
- Treatment group students were not expected to out-perform control students on **content area comparison** items (24 items).

The vocabulary part of the test consisted of items on the following three content areas: (a) Shoreline Science only, (b) Terrarium Habitat only, and (c) both Shoreline Science and Terrarium Habitat. The hypotheses were:

- Treatment group students were expected to out-perform control students on items covering **Shoreline Science only** (12 items).
- Treatment group students were not expected to out-perform control students on items covering **Terrarium Habitat only** (9 items).
- Treatment group students were not expected to out-perform control students on items covering **both Shoreline Science and Terrarium Habitat** (13 items), but it was possible.

Descriptive Results

As mentioned earlier, 25 teachers implemented the Shoreline Science materials in their classrooms, and another 10 teachers participated as control group teachers who used the GEMS materials. Twenty out of the 25 Shoreline Science treatment teachers submitted their data, and 9 out of the 10 control group teachers submitted their data. After excluding students and teachers (and schools) with missing values on the set of variables as specified for the Shoreline Science science data, we had a final sample of 400 students for the vocabulary analyses. For the analyses on comprehension scores, we further excluded an additional 11 classes/teachers because these teachers did not time the assessments as required. We had a final sample of 237 students.

Comprehension Data. This data set consisted of 237 students, with 192 students in the treatment group and 45 in the control group. Twelve teachers, at 12 different schools, taught the 192 treatment students, and three teachers at three different schools taught the control students. Table 4 has the detailed student distribution information. The distribution of male and female students in the data was similar to the distributions of them by group. The overall data consisted of 52% White students, 33% Hispanic students, and 15% students of the other ethnicities. White students were slightly over-represented, while Hispanic students and students of other ethnicities were slightly under-represented in the control group. In the control group, 69% of the students were White, 29% were Hispanic, and only 2% were of other ethnicities.

Table 4

Student Distribution by Student Level Variables for Shoreline Science Literacy Comprehension Data

Variable	Value Label	Treatment Group (*N*=192)	Control Group (*N*=45)	Overall (*N*=237)
Gender				
	Male	96	23	119
	Female	96	22	118
Ethnicity				
	All Other	34	1	35
	Hispanic	65	13	78
	White, Not Hispanic	93	31	124
Grade Level				
	Grade 2	71	38	109
	Grade 3	121	7	128

For the grade level, 46% of the students were in Grade 2 and 54% were in Grade 3. The distribution of grade level was very different for the control group. There were 84% Grade 2 and 16% Grade 3 students in the control group. In the treatment group, two-thirds of the students were third graders.

The average school enrollment was 452 students; the mean percent of students receiving free and reduced-fee lunch was 50%, and the mean percent of White students in each school was 53%. The 12 schools in the treatment had a slightly higher enrollment size (483 vs. 364), a similar percentage of students receiving free and reduced-fee lunch, and had 16% less White students in their schools than the 3 schools in the control group.

Vocabulary Data. This data set consisted of 400 students, with 274 students in the treatment group and 126 in the control group. Seventeen teachers, at 17 different schools, taught the 274 treatment students, and nine teachers at nine different schools taught the control students. Table 5 has the detailed student distribution information. The distribution of male and female students in the data was similar to the distributions of them by group. The overall data consisted of 59% White students, 29% Hispanic students, and 12% students of the other ethnicities.

Table 5

Student Distribution by Student Level Variables for Shoreline Science Literacy Comprehension Data

Variable	Value Label	Treatment Group (*N*=274)	Control Group (*N*=126)	Overall (*N*=400)
Gender				
	Male	136	64	200
	Female	138	62	200
Ethnicity				
	All Other	40	6	46
	Hispanic	79	38	117
	White, Not Hispanic	155	82	237
Grade Level				
	Grade 2	96	49	145
	Grade 3	178	77	255

White students were slightly over-represented, while students of other ethnicities were under-represented in the control group. In the control group, 65% of the students were White, 30% were Hispanic, and only 5% were of other ethnicities. For the grade level, 36% of the students were in Grade 2 and 64% were in Grade 3. The distribution of grade level was very different for the control group. There were 39% Grade 2 students and 61% Grade 3 students in the control group.

The average school enrollment was 447 students; the mean percent of students receiving free and reduced-fee lunch was 53%; and the mean percent of White students in each school was 68%. The 17 schools in the treatment had a slightly higher enrollment size (461 vs. 442), a higher percentage (5%) of students receiving free and reduced-fee lunch, and had 4% more White students in their schools than the 9 schools in the control group.

Using the available item scores, we created four test scores for comprehension and four test scores for vocabulary, for the pretest and posttest, separately. Please see Table 6 and Table 7 for students' mean test scores by their group membership and by item content category. These two tables also present students' scores in terms of gains, the gain scores between pretest and posttest. Across all three gain scores, treatment students had higher gain scores than control students. For the comprehension gain scores, treatment students out-performed control students by

about 1.5 points for all four sub-scores. For the vocabulary gain scores, the differences ranged from 1.5 points for Terrarium Habitat vocabulary only to 4 points for Shoreline Science and Terrarium Habitat vocabulary.

Table 6

Student Mean Scores on Shoreline Science Literacy Pre- and Posttests (Comprehension Tasks)

		Mean Test Scores		
Item Category	Variable	Treatment Group (*N*=192)	Control Group (*N*=45)	Overall (*N*=237)
Narrative				
	Pretest	4.88	3.96	4.70
	Posttest	7.72	5.36	7.27
	Gain	2.85	1.40	2.57
Target Science				
	Pretest	3.73	3.29	3.65
	Posttest	6.85	4.98	6.49
	Gain	3.12	1.69	2.85
Science Comparison				
	Pretest	3.82	3.44	3.75
	Posttest	7.02	4.78	6.59
	Gain	3.20	1.33	2.85
Content Area Comparison				
	Pretest	3.32	2.84	3.23
	Posttest	5.33	4.07	5.09
	Gain	2.02	1.22	1.86

Table 7

Student Mean Scores on Shoreline Science Literacy Pre- and Posttests (Vocabulary Tasks)

		Mean Test Scores		
Item Category	Variable	Treatment Group (*N*=274)	Control Group (*N*=126)	Overall (*N*=400)
Shoreline Science Only				
	Pretest	7.61	7.72	7.65
	Posttest	9.72	8.67	9.39
	Gain	2.11	0.95	1.74
Terrarium Habitat Only				
	Pretest	5.82	6.10	5.91
	Posttest	7.28	6.81	7.13
	Gain	1.46	0.71	1.22
Shoreline Science & Terrarium Habitat				
	Pretest	6.41	6.74	6.52
	Posttest	10.25	8.47	9.69
	Gain	3.84	1.73	3.18

HLM Results

The same estimation model was applied to analyze the four comprehension outcome variables and the three vocabulary outcome variables, as different hypotheses were made for these seven variables. Students' pretest scores on comprehension and pretest scores on vocabulary were used as control variables in the estimation equation, instead of pretest science MC and SA scores. Please see Table 8 and Table 9 for the HLM results.[11]

Results on Literacy Comprehension Gain Scores. As expected, treatment students out-performed control students significantly on the target science gain scores by 1.5 points, and treatment students also out-performed control students

[11] For these seven variables, the amount of variance found in the outcome variables that were caused by between-school differences varied from 2.5% (for Target Science gain scores, which indicates teachers/schools did not make any difference in the outcome variable) to 23% (for Shoreline Science & Terrarium Habitat gain scores, which indicates that about one quarter of the variation we found in student scores was due to teacher/school differences.).

significantly on the science comparison gain score by 1.6 points (see Table 8). The latter results indicated that students using the *Seeds/Roots* materials demonstrated their transfer ability in reading science texts in general. No statistical differences were found between students on narrative gain scores and content area comparison gain scores, as expected. The results also indicated that there were no differences between students of different ethnic groups, students of different grades, and students of different gender.

The results also indicated that students' pretest scores in comprehension were statistically significant related to their gain scores on narrative items (negative association), science comprehension items (positive association), and content area comparison items (negative association).

Table 8

HLM results Predicting *Shoreline Science* Literacy Gain Scores (Comprehension Tasks, *N*=237)

	Narrative		Target Science		Science Comparison		Content Area Comparison	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
School Level								
School average	1.06	0.92	0.29	0.94	-0.03	0.62	-0.08	0.56
Treatment	1.25	0.59	1.53*	0.27	1.63*	0.29	0.38	0.26
Student Level								
Pre-comprehension	-0.10*	0.03	-0.06	0.05	0.23*	0.28	-0.11*	0.03
Pre-vocabulary	0.07	0.04	0.09	0.05	0.63*	0.33	0.11*	0.03
Female	0.24	0.32	-0.05	0.37	0.41	0.46	0.37	0.45
Hispanic	0.12	0.60	0.79	0.73	-0.85	0.51	-0.11	0.33
Other	0.44	0.40	-0.52	0.45	-0.13	0.03	-0.77	0.48
Grade 3	0.36	0.38	-0.34	0.38	0.13	0.03	0.99	0.51

*p<.05.

Also, students' pretest scores in vocabulary were statistically significant and positively related to their gain scores on science comprehension items and content area comparison items. In summary, for students' gain scores on narrative and content area comparison items, their pretest scores on comprehension items had a negative association with the gain scores. Please note, though, the effect was 0.1—too small to make any substantial differences in practice. For target science gains

scores, none of the pretest scores on comprehension and vocabulary made a difference. For students' gain scores on science comparison, both pretest scores had a positive effect on the gain scores.

Results on Literacy Vocabulary Gain Scores. Table 9 reports the HLM results on the three literacy vocabulary gain scores. As expected, treatment students outperformed control students significantly on the Shoreline Science only gain scores by 1 point; and treatment students scored similar as control students on the Terrarium Habitat only gain scores. Unexpectedly, Shoreline only students also outperformed control students on Shoreline and Terrarium Habitat gain scores, by 2 points.

Table 9

HLM results Predicting Shoreline Science Literacy Gain Scores (Vocabulary Tasks, *N*=400)

	Shoreline Science only		Terrarium Habitat only		Shoreline Science & Terrarium Habitat	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*
School Level						
School average	4.05*	0.57	3.13*	0.45	4.98*	0.55
Treatment	1.01*	0.32	0.59	0.30	1.94*	0.43
Student Level						
Pre-comprehension	0.01	0.01	0.00	0.01	0.02*	0.01
Pre-vocabulary	-0.16*	0.02	-0.10*	0.02	-0.17*	0.02
Female	0.15	0.20	-0.01	0.19	0.01	0.22
Hispanic	-0.27	0.23	-0.32	0.27	-0.66*	0.27
Other	0.46	0.37	0.20	0.21	0.45	0.31
Grade 3	-0.18	0.40	-0.32	0.32	-0.23	0.46

*p<.05.

In terms of other student variables, there were no statistical differences between students of different ethnic groups, students of different grades, and students of different gender with one exception. Hispanic students were found to score two-thirds of a point lower than white students in the Shoreline Science and Terrarium Habitat vocabulary gain scores.

The results in Table 9 also indicated that students' pretest scores in comprehension were statistically significant related to their gain scores on Shoreline

Science and Terrarium Habitat vocabulary; a higher pretest score on comprehension was associated with a higher gain score. Students' pretest scores on comprehension were not statistically related to their gain scores on Shoreline Science and Terrarium Habitat vocabulary. And students' pretest scores in vocabulary were statistically significant and negatively related to their gain scores on all three vocabulary sub-test gain scores, approximately one-tenth of a point. However, the effects of both pretest comprehension and vocabulary scores were too relatively small to be of any substantive importance.

Quantitative Results for Terrarium Investigations Science

A total of 54 teachers participated in the *Terrarium Investigations* study. Twenty teachers implemented the *Terrarium Investigations* materials in their classrooms, and another 34 teachers participated in the study as control group teachers. All teachers submitted some materials for their students.

The results presented in this section compared treatment and control groups of students. Please see Appendix A for the more detailed results when the control group students and teachers were further divided into three sub-groups—Reader-only, Comparison, and GEMS.

Descriptive Results

After excluding students and teachers (and schools) with missing values on the list of variables used for the analyses, we had a final sample of 697 students taught by 46 teachers. Table 10 presents the student distribution information by whether they were in the treatment group or the control group. Out of the 697 students, 285 students were in the treatment group and the other 412 students were in the control group. The treatment students were taught by 17 teachers at 17 schools, and the control students were taught by 29 teachers at 29 schools.

Table 10

Student Mean Scores on Terrarium Investigations Science Pre- and Posttests

Item Category	Mean Test Scores		
	Treatment Group (*N*=285)	Control Group (*N*=412)	Overall (*N*=697)
Multiple Choice Items			
Pretest	2.54	2.26	2.38
Posttest	7.05	5.52	6.15
Short Answer Items			
Pretest	2.31	1.93	2.08
Posttest	6.37	4.85	5.47
Combined Items			
Pretest	4.85	4.20	4.46
Posttest	13.42	10.37	11.62

As shown in Table 11, the overall data consisted of 51% male students and 49% female students. Male students were slightly over-represented in the control group, at 54% and female students were slightly over-represented in the treatment group. In terms of students' ethnic makeup, there were more White students and fewer Hispanic students in the treatment group than in the control group, while the proportions of students of other ethnic groups were similar across these two groups. For the grade level, 40% of the students were in Grade 3, and the other 60% were in Grade 2. The distributions of grades were similar for both groups.

Table 11

Student Distribution by Student Level Variables for Terrarium Investigations Science Data

Variable	Value Label	Treatment Group (*N*=285)	Control Group (*N*=412)	Overall (*N*=697)
Gender				
	Male	134	223	357
	Female	151	189	340
Ethnicity				
	All Other	39	52	91
	Hispanic	23	97	120
	White, Not Hispanic	223	263	486
Grade Level				
	Grade 2	172	243	415
	Grade 3	113	169	282

The average school enrollment was 469 students; the mean percent of students receiving free and reduced-fee lunch was 39%; and the mean percent of White students in each school was 68%. The 17 schools in the treatment group had a lower enrollment size (464 vs. 472), a lower percentage of students (11%) receiving free and reduced-fee lunch, and 12% more White students in their schools than the 29 schools in the control group.

Students' mean test scores by their group membership and by item format were summarized in Table 12. Since students took part of the posttest as the pretest, we can not directly compare the scores to see how much they improved in the posttest. What we noticed was that control group students scored about a half point less than the treatment group students in the pretest. The difference increased to be about one point higher for the MC items, 1.5 points higher for the SA items, and 3 points for the combined items in the posttest.

HLM Results

The same estimation model for the Shoreline Science science analysis was applied to analyze the three outcome variables, except the outcome variables are student posttest scores. Students' pretest MC and SA scores were used as control

variables in the equation. Please see Table 12 for the detailed HLM results.[12] After controlling for student background variables and their pretest science scores, students who used the Terrarium Investigations materials scored about one point higher in the posttest MC and SA items than the control group students and the differences were found to be statistically significant.

We also found that students' pretest MC and SA scores were statistically significant for all three outcome variables. For the MC posttest scores, no student background variables were statistically significant predictors of student performance in the posttest. For the SA posttest scores, the results also indicated that female students scored higher than male students did (about one-third of a point), White students scored higher than all other students did (about half a point), and there were no grade level differences.

Table 12

HLM Results Predicting Terrarium Investigations Science Posttest Scores (*N*=697)

	Multiple Choice		Short Answer		Combined	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*
School Level						
School average	4.28*	0.32	3.05*	0.21	7.76*	0.62
Treatment	1.17*	0.18	1.12*	0.18	2.23*	0.35
Student Level						
Pre-multiple choice	0.29*	0.05	0.22*	0.04	0.50*	0.06
Pre-short answer	0.46*	0.05	0.65*	0.05	1.10*	0.10
Female	-0.03	0.11	0.32*	0.09	0.28	0.18
Hispanic	-0.36	0.22	-0.46*	0.15	-0.73*	0.22
Other	-0.34	0.25	-0.47*	0.13	-0.78*	0.35
Grade 3	-0.23	0.14	0.23	0.14	-0.02	0.21

*p<.05.

[12] For the MC and SA posttest scores, the amount of variance found in the outcome variables that were caused by between-school differences was 42% for MC and 41% for SA scores. The between-student differences made up for the other 58% (100% - 42%) of the variation found in MC scores and 59% in SA scores. The combination of explanatory variables accounted for 18% (for MC) and 26% (for SA) of the variation at the student level, and 58% (for MC) and 70% (for SA) at the school level.

Quantitative Results for Terrarium Investigations Literacy

There are a total of 54 teachers participated in the Terrarium Investigations study—20 teachers implemented the Terrarium Investigations materials in their classrooms and 34 teachers participated in the study as control group teachers. The research hypotheses were that no differences were expected in the comprehension items and some statistical differences were expected for the vocabulary items.

Descriptive Results

After excluding students and teachers (and schools) with missing values on the list of variables used in the analysis, we had a final sample of 556 students taught by 39 teachers. Table 13 presents the student distribution information by whether they were in the treatment group or control group. The 208 treatment students were taught by 13 teachers, and the 348 control students were taught by 26 teachers.

Table 13

Student Distribution by Student Level Variables for Terrarium Investigations Literacy Data

Variable	Value Label	Treatment Group (*N*=208)	Control Group (*N*=348)	Overall (*N*=556)
Gender				
	Male	97	188	285
	Female	111	160	271
Ethnicity				
	All Other	21	47	68
	Hispanic	17	90	107
	White, Not Hispanic	170	211	381
Grade Level				
	Grade 2	135	204	339
	Grade 3	73	144	217

As shown in Table 13, the overall data consisted of 51% male students and 49% female students. Male students were slightly over-represented in the control group at 54%, and female students were slightly over-represented in the treatment group at 53%. In terms of students' ethnic makeup, there were more White students and fewer non-White students in the treatment group than in the control group. For the

grade level, 39% of the students were in Grade 3, and the other 61% were in Grade 2. There were slightly more Grade 2 students and consequently fewer Grade 3 students in the treatment group than the control group.

The average school enrollment was 465 students, the mean percent of students receiving free and reduced-fee lunch was 32%, and the mean percent of White students in each school was 67%. The 13 treatment schools had a lower enrollment size (437 vs. 479), a lower percentage of students (9%) receiving free and reduced-fee lunch, and 19% more White students in their schools than the control group schools.

Students' mean test scores by their group membership and by item format were summarized in Table 14. We noticed that descriptively for the comprehension items in the pretest, treatment students scored lower in Danny and Soil items than control students, and they scored higher than control students in the same two categories in the posttest. No Beach items were included in the pretest. For the vocabulary items, treatment students scored higher in the pretest for both association and picture vocabulary items, and the differences became larger in the posttest respectively.

Table 14

Student Mean Scores on Terrarium Investigations Literacy Pre- and Posttests

		Mean Test Scores		
Item Category	Variable	Treatment Group (*N*=208)	Control Group (*N*=348)	Overall (*N*=556)
Comprehension Items				
Beach	Posttest	9.79	9.45	9.58
Danny	Pretest	5.83	7.69	7.00
	Posttest	9.56	8.86	9.12
Soil	Pretest	5.85	8.61	7.58
	Posttest	10.59	10.16	10.32
Vocabulary				
Association	Pretest	4.46	3.93	4.13
	Posttest	20.48	17.44	18.58
Picture	Pretest	2.50	2.36	2.41
	Posttest	9.56	9.06	9.25

HLM Results

The same estimation model was applied to analyze the three comprehension outcome variables and the two vocabulary outcome variables. Students' pretest scores on Danny on Soil were used as control variables for the Danny and Soil outcome variables, respectively. For the Beach outcome variable, we used students' pretest scores on both Danny and Soil items. Please see Tables 15 and 16 for the HLM results,[13] Table 15 having the results on comprehension posttest scores and Table 16 having the vocabulary scores results.

Results on Literacy Comprehension Scores. The treatment students scored similarly as the control students on all three areas of comprehension posttest scores. There were no statistical differences between these two groups of students. The HLM results also indicated that there were no differences between students of different ethnic groups, students of different grades, and students of different gender. Students' pretest scores in comprehension were found to be statistically significant related to their posttest scores.

Table 15

HLM Results Predicting Terrarium Investigations Literacy Posttest Scores (*N*=556)

	Beach		Danny		Soil	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*
School Level						
School average	3.63*	0.80	2.66*	0.71	4.60*	1.06
Treatment	1.85	1.19	1.97	1.01	1.75	1.56
Student Level						
Pretest	0.36*	0.03	0.75*	0.04	0.67*	0.05
Female	0.14	0.28	0.29	0.22	0.10	0.26
Hispanic	0.64	0.37	0.50	0.53	0.21	0.60
Other	0.13	0.39	0.55	0.32	-0.29	0.39
Grade 3	-0.54	0.53	0.26	0.42	0.27	0.40

*p<.05.

[13] For these five literacy posttest variables, their associated amounts of variance that were related to between-school differences varied from 14% (for Picture vocabulary scores) to 73% (for Soil scores, which indicates that more than two quarters of the variation we found in student scores were due to teacher/school differences.).

Results on Literacy Vocabulary Scores. The treatment students scored statistically higher than the control students on both association and picture vocabulary posttest scores. The difference was about 2 points for the association vocabulary scores and one-third of a point for the picture vocabulary scores. Students' pretest scores were a significant predictor of their posttest scores. For example, holding everything else constant, a student who scored one point higher in the pretest would score 1.4 points higher in the posttest than another student. The HLM results also indicated that there were no differences between students of different grades and students of different gender for association vocabulary scores. The same results could be applied to picture vocabulary scores with one exception - Hispanic students performed lower than White students.

Table 16

HLM Results Predicting Terrarium Investigations Literacy Vocabulary Scores (*N*=556)

	Association Vocabulary		Picture Vocabulary	
Variable	*B*	*SE*	*B*	*SE*
School Level				
School average	12.65*	0.76	8.18*	0.29
Treatment	2.10*	0.57	0.35*	0.11
Student Level				
Pretest	1.24*	0.13	0.40*	0.10
Female	0.45	0.41	0.09	0.09
Hispanic	-1.11	0.69	-0.43*	0.14
Other	-0.80	0.52	0.11	0.12
Grade 3	0.51	0.45	0.02	0.12

*p<.05.

Qualitative Evaluation

After obtaining the human subject approval from the University of California, Los Angeles' (UCLA) Office for the Protection of Research Subjects, the following email was sent to all Shoreline Science and Terrarium Investigations teachers asking for phone interview participants:

> CRESST/UCLA is looking for teachers who participated in *Seeds/Roots* units on Shoreline Science or Terrarium Investigations for a follow-up research study. You will be asked to schedule a one-hour telephone interview with us. During the interview, we will ask you questions about your classroom, your professional experience, *Seeds/Roots* unit implementation, and your opinion on the unit. The interview will be audiotaped with your permission for later transcription and coding of your responses. You will receive a stipend of $30 upon participation in our one-hour telephone interview.

Interested teachers were directed to contact us. The first seven *Shoreline Science* teachers and the first six Terrarium Investigations teachers who responded were recruited and scheduled for the one-hour phone interviews. The interviews address questions of x, y, z, and q. (see Appendix B for the interview protocol). The interviews were conducted in February 2005 for Shoreline Science and in March–May 2005 for Terrarium Investigations by an experienced researcher. All interviews were audiotaped and the tapes were sent to an outside agency for transcription. We used the interview dates and times to label the tapes and the subsequent interview documents. No teacher names were associated with these tapes and documents.

Shoreline Science Summary of Findings

The following list of bullet points describes our findings based on all seven interviews for each of the six research questions.

1. **How do the materials "work?" For example, to what extent and how are the units implemented? How engaged and motivated are students? What are teachers' reactions to the quality, usability, and utility of the units?**

   - All teachers used the Shoreline Science materials at least three times a week.
   - All teachers used the teacher's guide, the readers, the embedded assessments, magazine assessments, and pre- and post-assessments.
   - All teachers used the home activities, but not every teacher used all of them.
   - Most teachers used shared or paired reading in the classroom.
   - The teachers used the assessments to give grades, to check their students' understanding, to guide instruction, or to guide teaching.

- All teachers expressed that all their students took active parts throughout the unit and loved it, regardless of their previous achievement.

- All seven teachers found the curriculum/teacher's guide useful. It was clearly written and planned out, gave lots of ideas, strategies, and background knowledge, and laid out all the experiments.

- All teachers found the readers very useful.

- The assessments were very useful in measuring student growth, and most teachers thought they were useful for EL students too.

- All teachers thought the Unit gave well-balanced attention to science and literacy.

- All teachers said they would use the *Shoreline Science* Unit again if they had the choice.

2. What problems and/or misunderstandings do teachers and/or students encounter in implementing the materials?

- The time allocation was under-estimated.

- The pretests were too hard for the students, especially the literacy pretest.

3. What are the effects of using the materials on students' learning of science and reading of informational science texts?

- The Unit was very effective in helping students learn science and literacy, especially science.

4. For whom, for what purposes, and/or in what contexts do these materials appear most effective?

- For second graders, the Unit was most effective for those who were at or above grade level. For third graders, the Unit was effective for all students, especially the high achievers and those at or above intermediate language level.

5. What factors contribute to and/or detract from successful implementation? Some of the factors contributing to the successful implementation of the Unit were:

- student engagement

- the mix of activities and the sequence of the lessons
- use of group activities versus partner activities
- creation of a structure for students to internalize concepts
- the integration of literacy and science
- teacher enthusiasm and willingness to do the preparation work
- the building of both background knowledge and concepts
- connecting ideas in the books to actual experience

6. **How can the Units be improved? To facilitate implementation? To enhance students' learning in science and reading of informational texts? Some of the suggestions to improve the Unit were:**

    - need more ideas for the teachers on how to modify assignments for lower level students and special learners, especially for reading
    - highlight the must-do's or prioritize lessons to guide teachers with time constraints and guide teachers who use it as a main or as a supplemental program
    - have books of a lower reading level; for example: more pictures cards for support, more streamlined lessons that are not that long and not that many, and a few different vocabulary words each time instead of all at once
    - include more pictures for visual learners, especially for the teacher's guide
    - include more anecdotal information and background information in the teacher's guide

Terrarium Investigations Summary of Findings

The following list of bullet points describes our findings based on all six Terrarium Investigations teacher interviews for each of the six research questions:

1. **How do the materials "work?" For example, to what extent and how are the units implemented? How engaged and motivated are students? What are teachers' reactions to the quality, usability, and utility of the units?**

- All six teachers indicated they used them every day, except that one teacher skipped the instruction for four days due to a field trip.

- All teachers used the readers, the teacher's guide, embedded assessments, magazine assessments, pre- and post-assessments, and critical junctures.

- All teachers used the ELL considerations to a certain extent.

- Most teachers used all of the home activities. Two teachers did not use all of the home activities because of the weather.

- Independent, paired, and shared reading approaches were the three approaches that were used most often.

- Teachers used the assessments to evaluate student understanding, to guide instruction, and as part of the students' grades.

- All teachers said that their students took active parts throughout the whole unit. All students loved the Terrarium Investigations Unit.

- All six teachers found the curriculum/teacher's guide useful.

- In general all teachers found the readers very useful.

- Three teachers thought the unit gave balanced attention to science and literacy. The other three teachers would like to have more literacy contents.

- Five teachers stated that the assessments measured growth, one was not sure since she had not scored the posttests yet. One teacher specifically mentioned that the assessments were also useful to measure progress or mastery.

- All teachers would use the Terrarium Investigations Unit again if they had the choice.

2. **What problems and/or misunderstandings do teachers and/or students encounter in implementing the materials?**

   No problems reported.

3. **What are the effects of using the materials on students' learning of science and reading of informational science texts?**

- All teachers said that the students learned the literacy concepts and skills well.

- Most teachers believed the Unit challenged their students.

4. **For whom, for what purposes, and/or in what contexts do these materials appear most effective?**

   - The teachers thought the Unit was most effective for those who were at or above grade level and male students while being effective for everyone.

5. **What factors contribute to and/or detract from successful implementation? Some of the factors contributing to the successful implementation of the Unit were:**

   - student excitement and engagement

   - combine teaching the literacy development skills and having students apply the knowledge they gain through the hands-on activities

   - have interesting topics students could relate to

   - provide the needed materials to teachers

   - scaffold the concepts

6. **How can the units be improved? To facilitate implementation? To enhance students' learning in science and reading of informational texts? Some of the suggestions to improve the Unit were:**

   - include more support and suggestions for the less able ones in the teacher's guide

   - have different readers on the same topic for students of different levels

   - provide more support in terms of writing output and a bit more 'meat' in terms of vocabulary understanding and reading skill

   - have bigger books for shared reading

   - include examples of rubrics

   - have a training or demonstration video of how to use the unit

- separate the extra background information from the teacher's guide as the GEMS guides do

For the detailed interview information for the Shoreline Science unit, please go to Appendix C for the individual teacher profile and Appendix D for the summary of the seven individual teacher profiles. Appendix E and Appendix F have the corresponding individual and summary teacher profiles for the Terrarium Investigations.

Summary

CRESST evaluated the effectiveness of Seeds/Roots units Shoreline Science and Terrarium Investigations using both quantitative and qualitative approaches. Quantitatively, we analyzed the data collected during the study to document the Shoreline Science and Terrarium Investigations effects on student achievement, as compared to students who used other materials. Qualitatively, we interviewed a group of seven teachers who implemented Shoreline Science and six teachers who implemented Terrarium Investigations in their classrooms and solicited feedback on all aspects of the implementation. The findings from both approaches were positive and encouraging for both units.

Specifically for the study on the Shoreline Science unit, there were a total of 35 teachers participating: 25 in the treatment group using Shoreline Science materials and 10 teachers used non-Shoreline Science materials. By analyzing the collected data, we found that students whose teachers used Shoreline Science in the classrooms out-performed students whose teachers did not use Shoreline Science in all but one achievement measure where we had expected to see a difference. The differences were statistically significant for one of the two science measures, and for all four literacy measures that were expected. Compared to control students, Shoreline Science students statistically improved more in (a) overall science achievement, (b) science achievement tested in SA format, (c) literacy comprehension achievement in target science, (d) comprehension achievement in science comparison, (e) literacy achievement on Shoreline vocabulary, and (f) literacy achievement on Shoreline and terrarium vocabulary.

For the study on the Terrarium Investigations unit, there were a total of 54 teachers—20 in the treatment group using Terrarium Investigations materials and the other 34 teachers in the control group. Similar results could be concluded for the Terrarium Investigations unit. We found Terrarium Investigations to motivate

students to learn and to improve their learning in both science and literacy. The Terrarium Investigations students learned significantly more than the control group students in both science measures and both literacy vocabulary measures. Their posttest scores were statistically and significantly higher than the scores of the control students.

We also found that students' ethnicity and grade level had no effect on how much students learned from the materials once their pretest scores and treatment indicator at the classroom levels were controlled, as measured by the gain scores in the Shoreline Science unit. Our hypothesis is that the Shoreline Science assessment was a fair assessment, not biased against students of different ethnicities and students of different grades. The materials were equally effective in helping all students learn.

For the Terrarium Investigations unit, the findings were not so universal across measures. Ethnicity, gender, and grade variables did not have any effects on student posttest scores on science MC items, literacy association vocabulary, or on any of the three literacy comprehension measures once their pretest scores and treatment indicator at the classroom levels were controlled. For the science SA posttest scores, gender and ethnicity variables made some difference. For the literacy picture vocabulary posttest scores, Hispanic students scored lower that White students. These latter results, however, were consistent with what we would usually find in the literature. The majority of the Terrarium Investigations assessments were still found to be fair and unbiased against students for their ethnicity, gender, and grades.

The other interesting finding through analyzing the data was that Shoreline Science materials were effective in improving student learning and knowledge levels, especially for students without much prior knowledge. These students improved more than the other students when we looked at their gain scores between pre- and posttest scores. This could imply that Shoreline Science materials were effective in closing student achievement gaps.

As reflected in the teacher interviews, all seven Shoreline Science teachers and six Terrarium Investigations teachers reported that they highly valued the curriculum, the integration of science and literacy, and how the materials motivated and engaged their students. All teachers would use the materials again if there were no other requirements to fulfill. The Shoreline Science teachers indicated specifically

that both students and teachers learned the actual content knowledge, and teachers learned the teaching practice, teaching philosophy, and the realization how much 2nd and 3rd graders could learn with high-quality materials. Students started to enjoy science and they could hardly wait to learn things that were to be investigated in the classrooms. Teacher also reported improvement in their students' science and literacy knowledge.

In summary, by incorporating science and literacy into one set of curriculum materials, Shoreline Science and Terrarium Investigations were both beneficial to students and teachers. Students were motivated to learn, according to their teachers, and did learn significantly better than the control group students, even though all students improved their learning in both science and literacy. Teachers also were motivated to use the materials because the materials so strongly engaged students and because teachers found the materials useful for improving their own content knowledge and helping them learn new teaching activities and strategies. Both quantitative and qualitative approaches found Shoreline Science and Terrarium Investigations to be valuable instructional units.

References

Burstein, L. (1980). The analysis of multi-level data in educational research and evaluation. *Review of Research in Education, 4*, 158-233.

Cronbach, L. J., & Furby, L. (1970). How should we measure "change"—or should we? *Psychological Bulletin*, *74*, 68-80.

Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage Publications.

Lord, F. M. (1963). Elementary models for measuring change. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 21-38). Madison: University of Wisconsin Press.

May, K., & Hittner, J. B. (2003). On the relation between power and reliability of difference scores. *Perceptual & Motor Skills*, 97(3), 905-908.

Nicewander, W. A., & Price, J. M. (1983). Reliability of measurement and the power of statistical tests: some new results. *Psychological Bulletin*, *94*, 524-533.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Williams, R. H., & Zimmerman, D. W. (1996). Are simple gain scores obsolete? *Applied Psychological Measurement*, *20*(1), 59-69.

Appendix A
Terrarium Investigations Results on Four Groups

As described previously in the report, instead of having the students classified simply as treatment or control students, further distinctions were made regarding the control group students. There were three groups of control students—reader-only, comparison, and GEMS groups. The following results focus on how the results for these three groups of control students differed in the study.

Descriptive Results—Science

The student distribution information by gender, ethnicity, and grade level are presented in Table A1. Out of a total of 697 students, 285 students were in the treatment group, 117 in the reader-only group, 109 in the comparison group, and 190 in the GEMS group. Compared to the corresponding distributions for the treatment group, the comparison group had a larger percentage of male students. Comparison and GEMS groups had fewer white students and more students of all other ethnicities. The reader-only and GEMS groups consisted of more Grade 2 students; the comparison group had fewer Grade 2 students.

Table A1

Student Distribution by Student Level Variables for Terrarium Investigations—Science Data

Variable	Value Label	Treatment Group (N=281)	Reader Only Group (N=117)	Comparison Group (N=109)	GEMS Group (N=190)	Overall (N=697)
Gender						
	Male	137	59	63	98	357
	Female	144	58	46	92	340
Ethnicity						
	All Other	38	21	18	14	91
	Hispanic	37	12	22	49	120
	White, Not Hispanic	206	84	69	127	486
Grade Level						
	Grade 2	154	86	42	133	415
	Grade 3	127	31	67	57	282

Students' mean test scores by group membership and item format are summarized in Table A2. The treatment group had the highest pre-test and post-test MC (Multiple Choice) and SA (Short Answer) scores across all four groups of students. The GEMS students had the lowest pre-test scores and the comparison students had the lowest post-test scores.

Table A2

Student Mean Scores on Terrarium Investigations—Science Pre- and Posttests

Category	Variable	Mean Test Scores: Treatment Group (N=281)	Reader Only Group (N=117)	Comparison Group (N=109)	GEMS Group (N=190)	Overall (N=697)
Multiple Choice						
	Pretest	2.50	2.32	2.43	2.21	2.38
	Post-test	6.94	6.66	4.99	5.32	6.15
Short Answer						
	Pretest	2.29	1.97	2.12	1.83	2.38
	Posttest	6.35	5.32	4.81	4.63	6.15
Combined						
	Pretest	4.79	4.28	4.55	4.04	4.46
	Posttest	13.30	11.98	9.80	9.96	11.62

HLM (Hierarchical Linear Modeling) Results—Science

The HLM results on the science scores are shown in Table A3. For the HLM analyses, the treatment group was used as the base category—the results of the other three groups were compared to students in the treatment group. We found that compared to the treatment students:

- GEMS students scored significantly lower, by about one point on post-test MC and SA items, and about 2.5 points on the combined score;
- reader-only students scored about three/fourths of a point lower in the SA post-test items and scored similarly on the MC items; and
- comparison students scored significantly lower than the treatment students on both MC and SA items, as well as on the combined post-test scores.

These findings were observed when we held student background variables and their pre-test science scores constant in the estimation models. Students' pretest scores are always statistically correlated with students' posttest scores. Hispanic students scored lower than the White students on all three measures, and students

in the other-ethnicity category also scored lower than white students on both the SA and the combined post-test scores. Female students had a one-third of a point advantage over male students on the SA items.

Table A3

HLM Results Predicting Terrarium Investigations—Science Post-test Scores (*N*=697)

	Multiple Choice		Short Answer		Combined	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*
School Level						
School average	5.37*	0.23	4.18*	0.20	9.55*	0.36
GEMS	-1.28*	0.25	-1.26*	0.18	-2.54*	0.37
Reader Only	-0.14	0.22	-0.75*	0.32	-0.90	0.48
Comparison	-1.78*	0.29	-1.36*	0.34	-3.15*	0.52
Student Level						
Pre-multiple choice	0.29*	0.05	0.22*	0.04	0.51*	0.06
Pre-short answer	0.47*	0.05	0.65*	0.05	1.11*	0.09
Female	-0.04	0.11	0.32*	0.09	0.29	0.18
Hispanic	-0.44*	0.21	-0.52*	0.16	-0.91*	0.23
Other	-0.37	0.24	-0.49*	0.13	-0.84*	0.34
Grade 3	-0.22	0.14	0.21	0.16	0.00	0.23

*p<.05

Descriptive Results—Literacy

The final sample consisted of 556 students taught by 46 teachers. The treatment group had 207 students (13 teachers), the reader-only group 93 students (8 teachers), the comparison group had 117 students (8 teachers), and the GEMS group 139 students (10 teachers). See Table A4 for the student distribution by background variables.

Table A4

Student Distribution by Student Level Variables for Terrarium Investigations—Literacy Data

Variable	Value Label	Treatment Group (N=207)	Reader Only Group (N=93)	Comparison Group (N=117)	GEMS Group (N=139)	Overall (N=556)
Female						
	Male	101	47	66	71	285
	Female	106	46	51	68	271
Ethnicity						
	All Other	21	21	18	8	68
	Hispanic	30	12	25	40	107
	White, Not Hispanic	156	60	74	91	381
Grade Level						
	Grade 2	120	79	45	95	339
	Grade 3	87	14	72	44	217

As shown in Table A4, male students were slightly over-represented in the comparison group at 56%. In terms of students' ethnic makeup, the treatment group has the largest percentage of white students, the GEMS group the largest percentage of Hispanic students, and the reader-only group the largest percentage of non-White and non-Hispanic students. For the grade level, the distribution varies considerably across groups. The students in the reader-only group were mainly 2nd graders (85%), the comparison group had more 3rd graders (62%) and the treatment and GEMS groups were in between, but with somewhat more 2nd graders.

For the vocabulary tests, students' pretest scores were not very different from each other; students in the treatment group, however, seemed to do better on the

association posttest than students in the other three groups. For the comprehension items, students in the treatment group and the reader-only group had significantly lower pretest scores, but similar posttest scores as students in the comparison group and the GEMS group, on the subtests of *Danny* and *Soil*. The posttest scores for the *Beach* subtest were similar across the treatment and three control groups.

Table A5

Student Mean Scores on Terrarium Investigations—Literacy Pre- and Post-tests

		Mean Test Scores				
Category	Variable	Treatment Group (N=207)	Reader Only Group (N=93)	Comparison Group (N=117)	GEMS Group (N=139)	Overall (N=556)
Vocabulary						
Association						
	Pre-test	4.50	3.71	3.80	4.20	4.13
	Post-test	20.40	18.65	17.03	17.15	18.58
Picture						
	Pre-test	2.47	2.26	2.42	2.43	2.41
	Post-test	9.50	9.39	9.12	8.85	9.25
Comprehension Items						
Beach	Post-test	9.82	8.86	8.73	10.74	9.58
Danny						
	Pre-test	5.89	6.25	8.03	8.32	7.00
	Post-test	9.56	8.59	8.22	9.83	9.12
Soil						
	Pre-test	5.96	7.20	8.55	9.59	7.58
	Post-test	10.53	9.80	9.02	11.90	10.32

HLM Results—Literacy

The same estimation model was applied to analyze the three comprehension outcome variables and the two vocabulary outcome variables. The treatment group was used as the comparison group to examine the results of all the other groups. Students' pretest scores on *Danny* and *Soil* were used as control variables for the *Danny* and *Soil* outcome variables. For the *Beach* outcome variable, students' pretest

scores on both *Danny* and *Soil* items were used. See Tables A6 and A7 for the HLM results.[14] Table A6 shows the results on comprehension post-test scores, and Table A7 shows vocabulary score results.

Table A6

HLM Results Predicting *Terrarium Investigations* Literacy Post-test Scores (*N*=556)

	Beach		Danny		Soil	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*
School Level						
School average	5.45*	1.19	4.52*	1.19	6.16*	1.52
GEMS	-2.38	1.37	-2.48*	1.21	-2.49	1.71
Reader Only	-1.40	1.27	-1.08	1.14	-0.59	1.89
Comparison	-1.33	1.32	-1.62	1.05	-0.96	1.80
Student Level						
Pre-test	0.36*	0.03	0.75*	0.04	0.67*	0.05
Female	0.14	0.28	0.30	0.22	0.11	0.26
Hispanic	0.63	0.37	0.48	0.53	0.21	0.60
Other	0.11	0.40	0.51	0.33	-0.32	0.39
Grade 3	-0.60	0.56	0.23	0.47	0.25	0.43

*p<.05.

Results on Literacy Comprehension Scores. When student background variables and their pre-test science scores are controlled, the only significant difference we found is that GEMS students scored significantly lower than students in the treatment group on the *Danny* post-test items (about 2.5 points). There were no other significant differences for the other groups and on other sub-scores.

Results on Literacy Vocabulary Scores. As is shown in Table A7, GEMS students scored significantly lower, about 2 points, than students in the treatment group on the Association Vocabulary items. The difference between the treatment group and the GEMS group on the Picture Vocabulary items is not significant. Reader-only students scored similarly to the treatment students. The comparison

[14] For these five literacy post-test variables, their associated amounts of variance that were related to between-school differences varied from 14% (for Picture vocabulary scores) to 73% (for *Soil* scores, which indicates that more than two quarters of the variations we found in student scores were due to teacher/school differences).

group scored significantly lower than the treatment group—3 points lower on Association Vocabulary and about half a point lower on Picture Vocabulary.

Table A7

HLM Results Predicting *Terrarium Investigations* Literacy Vocabulary Scores (*N*=556)

	Association Vocabulary		Picture Vocabulary	
Variable	*B*	*SE*	*B*	*SE*
School Level				
School average	14.57*	0.85	8.47*	0.24
GEMS	-2.19*	0.84	-0.28	0.15
Reader Only	-0.49	0.71	-0.02	0.11
Comparison	-3.00*	0.67	-0.61*	0.19
Student Level				
Pre-test	1.25*	0.13	0.41*	0.09
Female	0.45	0.41	0.08	0.09
Hispanic	-1.25	0.67	-0.48*	0.14
Other	-0.90	0.50	0.08	0.12
Grade 3	0.70	0.48	0.08	0.11

*p<.05.

Appendix B
Teacher Interview Protocol

Teacher Interview Instructions

Thank you for agreeing to participate in this interview as part of CRESST/UCLA's evaluation of the Lawrence Hall of Science Seeds/Roots Program. The purpose of this interview is to gain a more detailed understanding of your experience with the Seeds/Roots' unit, either Shoreline or Terrarium, and its overall usefulness and effects on your instructional practice. We may also ask you some follow-up questions based on your responses to the end-of-unit survey. This interview will take approximately 60 minutes. We will pay you $30 for your time and for helping us.

Your name will not be disclosed in the final transcription of the interview, and any identifying information will be deleted from the final transcript. No information you provide as part of the interview will be linked to you or your school. The transcript is only available to the members of the UCLA Evaluation Team, and all contents of this interview are kept strictly confidential.

You may choose not to answer a question, and/or choose to terminate the interview if you do not feel comfortable. Participation in this study is strictly voluntary. The decision not to participate will in no way affect your relationship with the Seeds/Roots Program or with UCLA.

If there are sensitive issues that you would like to discuss, but prefer for them not to be entered into the transcription, the interviewer will honor your request and that portion of your interview will not be recorded or transcribed.

Do you have any questions so far? If at any time during the interview you don't understand a question, would like clarification, or would like the question repeated, please let me know.

If you are ready, let's start.

Teacher Background

1. What is your students' level of science proficiency? And how many percentages of them in each category?
   - Below Basic (Below grade level)
   - Basic (At grade level)
   - Advanced (above grade level)

2. What is your students' level of literacy proficiency? And how many percentages of them in each category?
   - Below Basic (Below grade level)
   - Basic (At grade level)
   - Advanced (above grade level)

3. What is your students' level of English language proficiency? And how many percentages of them in each category?
   - Below Basic (Below grade level)
   - Basic (At grade level)
   - Advanced (above grade level)

4. Do you have English learners (EL) in your classes?
   - If so, approximately what percentage of your students are ELs?
   - What's their level of proficiency in English language?
   - Do you give them special instruction?

5. Other than English learners, do you have other special needs students in your classes? (Please describe: learning disabilities, physical disabilities, other)
   - If so, approximately what percentage of your students are ELs?
   - Do you give them special instruction?

6. Have you participated in any other science professional development programs before? How many hours in the past 12 months? How many hours in the past three years?

7. Have you participated in any other literacy professional development programs before? How many hours in the past 12 months? How many hours in the past three years?

8. What teaching credentials or certifications do you currently hold? (e.g., clear credential in science/non-science subject, multiple subject elementary credential, CLAD/BCLAD certificate, or ELD certificate)

9. How would you rate your experience with using computers?
   Novice
   Low
   Moderate
   High
   Expert

10. How would you rate your expertise as a science teacher?
    Novice
    Low
    Moderate
    High
    Expert

11. Before teaching this unit, how knowledgeable were you about earth science?
    Novice
    Low
    Moderate
    High
    Expert

12. How would you rate your expertise as a literacy teacher?
    Novice
    Low
    Moderate
    High
    Expert

13. Have you ever participated in a GEMS workshop?
    Yes
    No
    Not Sure

14. What is your highest level of education?
    Associate's Degree
    Bachelor's Degree
    Certification beyond Bachelor's Degree
    Master's Degree
    Doctoral Degree

15. What was your major in college?

Experiences with Roots/Seeds

Evaluation questions: How do the materials "work"? E.g., to what extent and how are the units implemented?

1. Can you describe how you used the unit in your classroom?
    - *How often did you use the unit in your class? Everyday? A few days a week?*
    - *How many days did it take to complete the unit?*
    - *What's your time allocation in implementing the lessons and other suggested activities?*
    - *What materials did you use? (Teacher's guide, readers)*
    - *What activities did you use? (ELL considerations, home activities, assessments)*
    - *What reading approach did you use? (Independent reading, shared reading, paired reading, other) Why – the reasons to choose the particular one? How often?*
    - *Did you assign any home activities? Which ones?*
    - *Was the lesson adapted for different groups of students (ELs, basic, below basic)? If so, how?*
    - *Have you made any specific changes to your classroom or instructional practices as a result of the unit?*
    - *Did you use any other material to supplement the unit?*

2. How did you assess your students on this unit?
    - *Which type of assessment of the Seeds/Roots program did you use? (Pre-post assessments, Magazine assessments, Embedded tasks/rubrics, Critical junctures)*
    - *How did you use these assessments?*
    - *Do you think there were enough assessments available for you to identify the one that works best for you?*
    - *Were some assessments more useful than others? Please describe what and why.*
    - *Did you use any other assessments? Which ones? Why?*
    - *What rubrics did you use for scoring?*
    - *How useful were the rubrics in helping you score student work?*
    - *How did you use the information from the assessments?*
    - *What kinds of feedback did you provide to your students?*
    - *Were the assessments useful to measure progress? (ELs progress). Please explain.*
    - *How easy/difficult was the assessment system to use? (assessments, rubrics)Which ones in particular? Why?*

Effectiveness of Roots/Seeds

Evaluation questions: What are the effects of using the materials on students/ learning? Of science? Of reading of informational science texts?

For whom, for what purposes and/or in what contexts do these materials appear most effective?

What are teachers' reactions to the quality, usability and utility of the units?

1. Overall what do you think of the quality of the unit?
    - *Which part of the unit did you find most useful? Please explain.*
    - *Which part of the unit did you find least useful? Please explain.*
    - *Were there parts of the unit you would like to have seen covered more?*
    - *How useful was the curriculum guide? Why?*
    - *How useful was the teacher's guide? Why? Useful to guide instruction? Useful to use of books?*
    - *How useful was the reader? (reading level, science learning, literacy development)*

- *How useful were the ELL considerations? Which ones in particular? How well did they support language learning? How well did they support science learning?*
- *How useful were the home work activities?*

2. In your opinion, how effective was the unit?
   - *How well did your students learn the concepts and skills related to science? What concepts and skills were most difficult?*
   - *How well did your students learn the concepts and skills related to reading? (Learning and practicing reading strategies). Please describe the most important learning achievements.*
   - *How well does the unit meet your students' interests?*

3. For what groups of students was the unit most effective?
   - *How well worked the unit for EL, basic, below basic, advanced. Please explain.*
   - *How, if at all, do you feel the program could be changed to better meet their needs?*

Implementation of Roots/Seeds

Evaluation questions: What factors contribute to and/or detract from successful implementation?

How engaged and motivated are students?

How can the units be improved? To facilitate implementation? To enhance students' learning in science and reading of informational texts?

What problems and/or misunderstandings do teachers and/or students encounter in implementing the materials? Are teachers able to understand the materials and how to implement them? Are students able to understand directions, accomplish tasks as intended?

1. What was your level of ease/difficulty in implementing the unit?
   - *What was easier for you to implement? Why?*
   - *What was the most difficult? Why?*

2. What, if any, problems or misunderstandings did you encounter during the implementation?

3. Overall, how did your students react to the unit?
   - *Were the students actively taking part in the exercises? Please give examples?*
   - *How, if so, were the students being challenged by the content (science, reading)?*
   - *What were the students' reactions to the unit? What part of the unit do they like and dislike?*
   - *Did students need support in using the materials?*
   - *What problems, if any, did your students have understanding directions and accomplishing the tasks as you intended?*
   - *What can we do to make the unit more useful to the students?*

4. How well did the unit balance attention to science and literacy? Please explain.

5. Would you use the unit again? Why? How?
   - *Main science/literary program, Supplement to existing program, Other?*
   - *What parts in particular (Activities, assessments)*

6. Are there any changes you would recommend for the unit?
   - *Improvement of any materials, guidance, content for students' learning (science, reading), implementation?*

7. What, if anything, did you learn from the Seeds/Roots unit?
   - *How do you feel that the unit prepared you to teach your students?*
   - *How, if at all, has the unit led you to change your instruction to better support your students?*

8. In your opinion, what do you think contributes to the success of implementation of the unit?

Is there anything you would like to add about your experiences with Seeds/Roots? If not, thank you for your time and participation.

Appendix C
Shoreline Science Teacher Profiles

Each of the seven individual interview responses were categorized and summarized in order to answer the evaluation questions listed on page 4. Besides providing answers to the specific evaluation questions, the teachers also provided some of their background information in teaching, professional knowledge, their classrooms, and their students. This information is presented at the beginning of each teacher profile under "Classroom Background" and "Teacher Background." Since some of the teachers we interviewed also provided information on how *Shoreline Science* had changed their instructional practices and improved their learning, we present that information, when available, as item 7 in each profile. These seven individual profiles are presented in this Appendix for a quick insight into each teacher's specific responses during the interview.

Profile for Teacher 1

Classroom background. This teacher used the *Shoreline Science* Unit with her second-grade students. Seventy percent of the 18 students were estimated to be proficient and 30% not proficient in their science and literacy knowledge. There were three ELLs below the level of advanced intermediate, and one student was a special needs student (resource student).

Teacher background. This teacher spent 8 hours in the past 12 months and about 15 hours in the past 3 years in science professional development. She also attended 20-30 hours of literacy professional development in the past 12 months and about 60 hours in the past 3 years. She has a BA degree, a California Preliminary (5 years), and has the Cross-Cultural, Language and Academic Development (CLAD) teaching credential. She ranked herself as "moderate to high" in her computer experience/knowledge, and "moderate" in her Earth science knowledge, as a science teacher, and as a literacy teacher.

Unit implementation. The Unit took the teacher about 10 weeks to complete. She used the *Shoreline Science* materials 3 or 4 days per week for an hour a day at the beginning, and used them every day at the end to finish and to meet the deadline. She did all home activities, all assessments booklets, and some of the English language learner's materials with the whole class because she perceived that all second graders were English language learners to some extent. This teacher followed everything in the Unit except for the scoring rubrics. Instead of using the suggested 4-point scoring scale, the teacher used her school's 5-point grading system for all assessments besides the pre- and posttest due to her school requirement.

She used *Shoreline Science* for science and Houghton Mifflin for literacy. The ELLs got extra help in interpretation of words. In the classroom, the teacher typically had the whole class reading before paired reading as she found this practice prompted students to ask questions. She used the information from assessments for the purposes of grading and identifying students who were weak in certain areas.

Student engagement and motivation. *Shoreline Science* got students really interested in learning. They loved reading the magazines and loved the hands-on experience. When asked to write about school in January 2005, almost all of the students wrote that their favorite subject was science. The teacher also heard her students explain how interesting and fun the Unit was to a new student.

Unit quality, usability, and utility. The teacher found all the teaching materials useful in giving her content information and teaching ideas, preparing the experiments, measuring student progress, and providing students with proper text information. The clear layout also made it easy for the teacher to implement the Unit and easy for students to follow and understand. The timeline was helpful in keeping the teacher on task. The four homework activities were great and the students loved them. The Unit was quite balanced between science and literacy and all of the science activities had literacy involved. The teacher will use the Unit as a supplement since she has to use the state adopted curriculum.

Problems in implementation. Assessment was the most difficult part of the Unit. It was confusing to determine which were regular assignments and which were assessments, and it was hard to use the rubrics. The teacher would like to have more ideas on how to modify assignments for slower students. The Unit also took more time than expected.

The second graders had a hard time with the language arts test, especially the Mayes Close Comprehension. They could hardly get any points on that because they were second graders and they were not fast readers. The students felt really stupid, and some cried during the pretest. In terms of materials, *Gary's Sand Journal,* Activity Four, and some Readers were too hard for second graders to understand, especially for the slower students. In terms of activities, the Discourse Circles were too much for the second graders. They could not grasp the concepts and did not understand what they were supposed to be talking about.

Student learning. Students loved the whole Unit, especially Activity Three, *Shoreline Science Organism*. They learned how to use the index and the glossary in the book, how to find information, how to stay on topic, and how to write a report. So they learned a lot in both literacy and science. The students were challenged by the contents and impressed by how much they learned and how much it expanded their thinking.

Most effective for. The Unit was most effective for the proficient and advanced second-grade students.

Factors for successful implementation. There were three main factors that contributed to the successful implementation: (a) the engagement and interest of the students; (b) the teacher's enjoyment doing the activities; and (c) the teacher's motivation when seeing the students excited about learning.

Recommendations and suggestions.

- need wider line space for the worksheets
- more support in helping slower students in Grade 2 classrooms
- more content materials on the tides part
- more literacy things like fluency training or taking apart big words, etc.
- have directions for the experiment right in that lesson to avoid flipping back and forth
- preference for the magazine assessments to be classroom activities so she could talk about it in depth with her students
- not having the pretest because it made students feel bad about not knowing the content before instruction

Other comments on instructional changes and teacher learning.

- The teacher used more science vocabulary in her teaching because students did learn the vocabulary and liked having their science words versus everyday ones.
- Because students were learning with investigations, the teacher now let them investigate before giving the lesson instead of doing the lesson and then doing the experiment.
- She learned the science content areas that were covered and learned how to do vocabulary with second graders.

Profile for Teacher 2

Classroom background. This teacher used the *Shoreline Science* Unit with her second-grade students. Thirty percent of the 26 students were estimated to be at grade level, 20% above grade level, and the rest below grade level in their science and literacy knowledge. Out of the 26 students, two spoke English as a second language and two were in a special education program. Three of these students were significantly below grade level and one was on grade level. The teacher would work with them in a small group on more specific skills and/or had volunteers work with the special needs students individually.

Teacher background. This teacher did not participate in any science professional development in the past 12 months or the past 3 years; the teacher attended zero hours of literacy professional development in the past 12 months and about 7 hours in the past 3 years. She has a master's degree in education and holds an Illinois Standard Elementary Teaching Certificate. She ranked herself as "moderate" in her computer experience/knowledge, in her Earth science knowledge, and as a science teacher; she ranked herself as "high" as a literacy teacher.

Unit implementation. The teacher did not finish the Unit in the allocated time period, but she wrapped up the Unit as of February 15, 2005. During the period from September 2004 to February 2005, she used the Unit three days a week on average, an hour to an hour and a half each time for the first three sections and two and a half hours for the last section.

She followed the curriculum explicitly, did all assessments except the critical junctures and the second magazine assessment due to lack of time, and used all materials except that she only used one homework activity (writing a letter about getting a sample). She supplemented the Unit with some journal writing as a daily activity. She brought in books on animals and organisms from the library, and she added some activities related to the vocabulary. In the classroom, the teacher would start by having the children read the book with a partner, had them talk about it, and then read it again the next day or however the curriculum advised it be done. Then toward the end, the teacher would read to the children before they did the independent reading.

The teacher used the assessments to measure students' understanding and application and to evaluate their performance for the grades with respect to the

rubrics. The assessments were useful for her to measure students' progress. She used multiple scoring rubrics and discussed the rubrics extensively with her students as well as posting them on the board to make students aware what they needed to demonstrate for each score point.

Student engagement and motivation. The children responded to the books very well. The readers held their interest even though they were very challenging. The students loved the Unit and really enjoyed it. Students liked exploring their sand; they liked looking at the sizes and the shapes; they enjoyed the journal; they liked *Sandy's Journey to the Sea*; they liked studying about their organisms; and they were very proud of their reports. There was not one bored person in the class. The Unit came alive for them.

Unit quality, usability, and utility. The teacher viewed the quality of the Unit to be very, very good—it was well thought out, had a good scope and sequence of the lessons, was very thorough, moved at the right speed, and gave students a chance to discover for themselves the concepts of the models that were built into the activities. The lessons were very meaningful and interesting for the students, especially the lessons on exploring the beach and the oil spills. The teacher mentioned that it was awesome the way it sequentially presented science concepts and involved the children in hands-on discovery activities and inquiry-based learning. The teacher's guide was an excellent resource and well-planned and guided her instruction.

The teacher found the concepts in the lessons on *Where Sand Comes From* easy and the lesson on organisms a challenge for her students. She really liked the report about the organism though. In order to do the report they had to understand habitats on the sand, under the sand, and in the nearshore waters; they had to understand other organisms in order to discuss predator and prey; and they had to understand the structure of their organisms and the behaviors of their organisms. The report required student skills in language, writing, reading, organizing, and content knowledge. The teacher found the activity on making concept maps useful for student understanding of concepts and in their report writing.

She thought *Shoreline Science* was balanced between science and literacy, though in the future she would use *Shoreline Science* as her main science curriculum and as a supplement for her literacy curriculum.

Problems in implementation. It was not totally realistic to cover some of those lessons in 60 minutes. Some of them took two periods. The assessment system was very time-consuming to use. It was too difficult to have three children to a book.

The Unit was very challenging even for her high-performing readers. However, going through the books twice and then referring to the books would be enough for the children to internalize the terms and the concepts that were presented. The sand journal was a little bit long. The teacher started to lose students. She also thought the students had a hard time working in groups of four, because group dynamics for second graders are really hard. She preferred groups of two instead of four students.

Student learning. Her students learned a lot, even her lower students. For instance, they used the terminology in their report and on their posters. It was very clear that they understood what they had written on the posters and what she had written in their reports. The teacher was very pleased with what they learned, with what she thinks they know, and with what they're taking away with them. Her students at and above grade level definitely understood the Unit very well. The teacher felt her students could teach it.

Most effective for. The Unit was most effective for high-achieving second-grade students.

Factors for successful implementation. The mixture of activities, the lesson sequencing, and the group-versus-partner activities contributed to the success of implementation of the Unit.

Recommendations and suggestions.

- have a reader for every two students
- need more attention to vocabulary development
- maybe put the name of the organism in bold print or have more pictures
- maybe shorten some of those discussion lessons and/or include some other activities that would engage every learner
- more lessons on writing
- the Sand Journal was too long. It could be shortened.

- need more time for the Unit

Other comments on instructional changes and teacher learning

- The teacher now used a glossary on the wall.
- She used the shared partner activities.
- She used concept maps.
- She included some the report-writing techniques in her teaching.

Profile for Teacher 3

Classroom background. This teacher used the *Shoreline Science* Unit with her second-grade students. Twelve out of the 17 students were estimated to be at grade level, and 5 to be below grade level in their science knowledge. Ten students were estimated to be on grade level and 7 below grade level in their literacy knowledge. There were 2 ESL students who were below grade level who received special instruction outside the classroom. Three students were being tested as special needs students.

Teacher background. This teacher spent 40 hours participating in science professional development in the past 12 months and 150 hours in the past 3 years; she attended 20 hours of literacy professional development in the past 12 months and 75 hours in the past 3 years. She has been teaching for about 10 years, holds an elementary self-contained teaching credential (gifted and talented) and a BS degree. She ranked herself as "pretty high" in her computer experience/knowledge and as a science teacher; "high" in her Earth science knowledge; and "moderate" as a literacy teacher. The teacher has been using GEMS for the past 7 to 8 years.

Unit implementation. The teacher taught the Unit every day since September except the few weeks she had to work on other materials. However, as of January 15, 2005, she has not finished all four activities. She stopped at Lesson Three in Activity Four. She stated that it took her two to four times more time than was specified to finish, and sometimes she would spend 4 hours a day using the Unit. She did all home activities, all assessments booklets, and used some of the English language learner's materials for her slower students.[15] Both the assessments and their rubrics were pretty easy to use, but she did revise some of the assessments so her students could understand them better.

She used *Shoreline Science* for both science and literacy. In the classroom, the teacher typically paired students up and had them read the book along with her on the first day and had students take turns reading parts the next day. Sometimes she would also follow up by having students work independently. She used the information from the assessments for the purposes of grading, identifying students who needed extra help, and improving her own teaching practice. The teacher

[15] Home activities were done in the classroom instead of at home. The teacher later reflected that if she had her students do the home activities at home, the activities could have given parents an idea what was going on in the classroom, thereby enhancing classroom learning.

supplemented the Unit with books from the library so the students had a larger variety of books. She also elaborated on some of the activities by incorporating technology.

Student engagement and motivation. Since the students were very interested in the materials and activities, they thought science was the best thing in the whole wide world. The teacher used library books to supplement the Unit because students were so into learning. She also heard her students talking, and using the scientific terms they learned months ago.

Unit quality, usability, and utility. The teacher found all the teaching materials useful, especially the writing component. She especially liked *Sandy's Journey to the Sea* because it was such a meaningful and interesting activity. Students still remembered it a few months later. Students really liked the *Jolly Rancher* activity on how sand is made; the globe activities where they tried to go around the globe with their finger over water and then over land; and the oil spill. She was able to elicit information from her students in ways she was not able to do previously. In her opinion, the Unit was balanced between science and literacy. This teacher would use *Shoreline Science* as her main program.

Problems in implementation. The teacher did not have any problem in implementing the Unit. Related to her level of comfort with science and writing, she found the science activities easy to implement and the writing the most difficult to implement. Sometimes the students needed a little bit of support so she modeled a few things for them.

Student learning. Students liked the whole Unit. There was no single part they did not like. It was very effective in "pulling" information from the students. Students learned how to write a report, and she was very impressed with the quality and ability of the writing her students produced. The Unit made students think in different ways, and that was the biggest challenge for them since they were pretty much used to seeing worksheets.

They learned how to use the index and the glossary in the book, how to find information, how to stay on topic, and how to write a report. So they learned a lot of literacy and science at the same time. The students were challenged by the contents and impressed by how much they learned and how much it expanded their thinking. They learned how to make inferences.

Most effective for. The Unit was most effective for second-grade students at or above grade level.

Factors for successful implementation. The teacher attributed the success to the fact that literacy, reading, writing, science, and experiments were all integrated in the *Shoreline Science* material. Students got constant reinforcement on what they'd read about, what they'd done, and then what they had written about. They got to move around and do things. It was visual, oral, and physical.

Recommendations and suggestions.

- more content materials (background) for teachers, especially on organisms
- need to know how to condense the Unit or shorten it somehow if there is a lack of time

Other comments on instructional changes and teacher learning. The teacher incorporated techniques she learned from the Unit for teaching literacy and writing, etc.

The Unit helped her to be more creative in building other curriculum topics (thinking outside the box); improved her confidence level with writing; and taught her a lot of content knowledge about beaches, organisms, etc.

Profile for Teacher 4

Classroom background. This teacher used the *Shoreline Science* Unit with her third-grade students. Ten to fifteen percent of the students were estimated to be at grade level, and the rest below grade level in science knowledge. One hundred percent of the students were considered below grade level in literacy knowledge. Out of the 20 students, 2 were at the beginner level, 5 at the early intermediate level, and the other 9 at the intermediate level. All of them got special instruction. This teacher did not have any other special needs students.

Teacher background. This teacher spent 3 hours in science professional development in the past 12 months and 10-15 hours in the past 3 years; and she attended 30-40 hours of literacy professional development in the past 12 months and about 100 hours in the past 3 years. She has a BA degree and holds the California multiple subjects for K-5, the national certification, a BCLAD, and a supplemental authorization to teach Spanish through 9th or 10th grade. She ranked herself as "moderate" in her computer experience/knowledge, in her Earth science knowledge, and as a science teacher; she ranked herself as "experienced" as a literacy teacher.

Unit implementation. The Unit took the teacher about 8 to 9 weeks to complete, a bit longer because she had to break some lessons into two. She used the *Shoreline Science* materials every day for one to one and a half hours. She used everything that was provided except that she only used two of the home activities—*Your Favorite Beach* and *Investigating Oil,* and she translated both into Spanish for her students due to their low level of English language proficiency.

The teacher used some of the suggested ELL considerations and adopted preview/review practice in her instruction for her ELLs, besides delivering the instruction in both English and Spanish. She also brought in newspaper articles that were related to oil spills for her students. She used the scoring rubrics that were provided. Typically she used assessment information to drive instruction, such as in the areas that they're falling out or needed more support in; but for *Shoreline Science* she did not go back and re-teach due to time constraints. The teacher used a variety of reading approaches. Most of the time her students did shared and paired reading first (since half of them could not read the materials independently), some guided reading, and then independent reading for the higher performing students.

Student engagement and motivation. Students found the Unit interesting, engaging, and motivating. They were so excited to be doing hands-on lessons that they loved. They would frequently make comments like "Wow, this is the best class ever. We never got to do anything like this before!" They would cheer when they started the lessons, got really involved in the lessons, and really responded well to the lessons. They loved the Unit.

Unit quality, usability, and utility. The teacher liked *Shoreline Science* very much. She liked how it was divided into before reading activities, during reading activities, and then post reading activities. She also liked the think/pair/share routine and parallel line/partner share activities that were good for working with English learners to lower the anxiety and keep everybody active. The hands-on activities and readers were very useful. The readers were excellent in terms of linking with the hands-on and bringing it back into the text and reinforcing concepts, and they were right to the point, very readable, totally applicable, and had really good illustrations and photos. The readers were very readable, even for the intermediate English learner.

Having the estimated time helped the teacher to stay on track, even though sometimes it went a lot longer than the estimated time. She thought the writing activities were pretty good. She found the rubrics useful as they emphasize understanding, but they took more time to score. The teacher thought the pre- and posttests were very useful in measuring student progress. In her opinion, the Unit had a pretty good balance between science and literacy. In the future she would use *Shoreline Science* as a supplemental program to both science and literacy by using the concepts but not going into as much depth.

Problems in implementation. The students had difficulty in doing the following: (a) word search because they were not familiar with the format and did not have enough time; (b) crossword puzzle because her students did not get the given clues; and (c) discourse discussions because of students' limited language skills.

The teacher found it hard to implement the Unit because of the time factor—she had to do Open Court reading, and some lessons took longer than the estimates. She found the assessments and the rubrics very time-consuming.

Student learning. Students gained a tremendous amount of science knowledge, and a lot of new vocabulary after seeing the words in multiple contexts.

They also learned to apply the knowledge in later new contexts. For example, during a field trip to the Second Nature Center in Long Beach, students used terminology like "adaptation" in their conversations, which really impressed the tour guide. They also learned to write reports using the index and the table of contents, etc.

Most effective for. ELL students at the intermediate or higher language levels in third grade.

Factors for successful implementation. Teacher enthusiasm and willingness to do the work.

Recommendations and suggestions.

- need a Spanish version for schools in California
- having books of lower reading level and having a smaller number of vocabularies to work with for new arrival students that do not speak English
- include more pictures for support
- mark activities/lessons as must-do's and optional for people with time constraints
- exclude the class report lesson in which students copy paragraphs they wrote earlier

Other comments on instructional changes and teacher learning. The teacher incorporated the concept wall into her regular lessons as a result of the Unit.

Profile for Teacher 5

Classroom background. This teacher used the *Shoreline Science* Unit with her third-grade students. One hundred percent of her 27 students were estimated to be below grade level in science, and for literacy knowledge, 75% were below grade level and 25% at grade level.[16] Five students received special education. There was no special instruction in the classroom. Instead, the teacher would do small group reading with her below-grade-level students so she could focus on certain groups of students who needed help. Sometimes she would have other teachers come in to work with a small group of students. The teacher had one ELL student below grade level who would leave the classroom for some special instruction.

Teacher background. This teacher spent 25 hours in science professional development in the past 12 months and about 100 hours in the past 3 years; and she attended 10 hours of literacy professional development in the past 12 months and about 100 hours in the past 3 years. She has two master's degrees, one in education and one in business. The teacher holds a certification in K-6, called Common Branches in New York State. She ranked herself as "moderate to high" in her computer experience and as a science teacher, as "novice" in Earth science, and "high" as a literacy teacher.

Unit implementation. The teacher had only covered two lessons out of the *Shoreline Science* Unit at the time of the interview because she had to teach some other required curriculum. She started the Unit in late September. The teacher was planning to finish it. She used the *Shoreline Science* materials four times per week for one hour each time. Even though she tried to finish the materials in the suggested time, sometimes she broke lessons into two for her low-proficiency students. She used everything that was provided, though she revised the assessments provided, added in some of her own home activities, did not do the postcard home activity, and she brought in supplemental materials related to *Shoreline Science* to enhance students' low prior knowledge. She did not use the ELL considerations in exact form, but they influenced her teaching. The teacher used the assessments to see what kids understood from the instruction.

The teacher typically started with shared reading, buddy reading, or partnership reading and then had the kids re-read the books independently in the

[16] She started with 21 students, and got some new students in the middle of the Unit.

independent reading workshop while she worked with a small group. She used *Shoreline Science* as a supplement to her science and literacy textbooks from Houghton Mifflin.

Student engagement and motivation. The students loved the Unit and loved using the materials. They liked *Sandy's Journey to the Sea*. They complained a bit about the writing, but they liked to write the postcard and then send it. Some of them liked the rock boxes; it was as if they had not seen anything like that before. They really liked creating the website, drawing Sandy's journey, writing postcards, etc., and their favorite was the Jolly Rancher Activity.

Unit quality, usability, and utility. The teacher found the guide and readers to be especially imperative, because they provided lots of information, were well organized, had great pictures that helped her understand the materials, and (readers) enabled her to identify her students' proficiency level in both reading and understanding science. She liked the way discourse circles were organized, and the think/pair/share activity routine. She liked the activity of building a knowledge (concept) map and using the web. It was very visual and meaningful to her kids.

She found the amount of assessments to be sufficient. She liked one particular component (close activity?) in the pre- and posttests instructional helpful. The activity enabled her to see which kids were re-reading and which kids were self-monitoring and cross-checking, to see which word was the best word, to assess how quickly kids could read, and how fluent they were. She really liked the assessment at the end of the first lesson because it assessed student understanding, their writing skill, and whether they used the vocabulary from the Unit. She did not like the paragraph assessment at the end of Unit two because she felt it was not exactly a clear assessment of what they knew.

She found it amazing the way *Shoreline Science* built concepts and how the literacy and science were combined while still being accessible for third graders. The Unit provided a variety of things to do, and made kids think and share their thinking with one another. She thought the Unit was well balanced between science and literacy. She would use the Unit as her main text for science.

Problems in implementation. There was not enough time sometimes, such as for the creation of the knowledge (concept) map and the activity on what sand was composed of. These took more than one lesson, and could easily take three or four lessons to build the concepts. The teacher had to slow down a bit because of the

profile of her readers. It took her some time to figure out how to apply the rubrics. The teacher would also like to have more information on the Jolly Rancher materials.

The discourse circles were very challenging for her students as they were not as articulate, and they did not really reason that well. It was difficult for students to understand minerals because they did not know what they were.

Student learning. All students gained lots of knowledge from the Unit, even the extremely low-performing kids. They learned both content knowledge and learning skills, like how to make knowledge (concept) maps, how to use the web to help with their writing, what a Shoreline is, and how to make sand over time.

Most effective for. The Unit was most effective for her third-grade high achievers and students who were more curious about things and were able to stick with it.

Factors for successful implementation. There were two main factors that contributed to the successful implementation: (a) teacher's desire to implement the Unit; and (b) the framework underlining the Unit—building both knowledge and concepts. Connecting the content in the readers with the inquiry experiences was a powerful learning combination for students.

Recommendations and suggestions.

- Overview in teacher's guide and lessons sometimes gave different information. Maybe the information in the teacher's overview could be connected with the lessons for easier access.
- need more background information for the students in some of the materials, like minerals
- more pictures would be helpful
- prefer to have comments on what are the must do's and what are secondary things, especially for first time users
- maybe the first two lessons could be expanded while shortening the rest since the Unit took a lot of time

Other comments on instructional changes and teacher learning.

- She learned how to create a knowledge (concept) map.

Profile for Teacher 6

Classroom background. This teacher used the *Shoreline Science* Unit with her third-grade students. Ninety-nine percent of her 25 students were estimated to be below grade level in science and 1% at grade level. For literacy knowledge, 10% were below grade level, 80% at grade level, and 10% above grade level. There were no ELL students in her class. She had four language resource students. The resource students were paired up with other students for buddy reading when the other students did independent reading, and their writing assignments were shortened.

Teacher background. This teacher spent 3 hours in science professional development in the past 12 months and about 300 hours in the past 3 years; the teacher attended 10 hours of literacy professional development in the past 12 months and about 60 hours in the past 3 years. She has a BA and holds a K-8 standard teaching certificate and a provisional gifted endorsement. She ranked herself as "low" in her computer experience, and "high" as a science teacher, in Earth science, and as a literacy teacher.

Unit implementation. The teacher used the *Shoreline Science* materials four times a week for an hour and a half each time. She finished the *Shoreline Science* Unit in about 50 days instead of the 40 days suggested. Other than that she only did the first home activity due to lack of parental support, the teacher used all the materials that were given and followed them faithfully. She did not use any other materials to supplement the Unit. She also did not use any of the ELL considerations.

The teacher did not use assessments for assigning grades. She used them to identify students' level of understanding. She did not use the information to adjust her teaching on *Shoreline Science* materials, to follow the flow of the Unit, or to influence the results of the study. The teacher used a combination of reading approaches with shared reading as the one most used.

Student engagement and motivation. *Shoreline Science* was of great interest to the students, especially since there were no oceans around, and the ocean was a foreign concept to the students. They were very excited about learning it every day. They could not wait and always asked, "What are we doing in science today?" Every day they were ready to tackle it and could not wait to see what they were doing next because they were interested.

Unit quality, usability, and utility. The teacher found *Shoreline Science* to be a great resource for providing a lot of background knowledge. She especially liked the format of covering the science content before moving to literacy, bouncing between these two during the lesson and then having the readers backing up both the science content and literacy. She thought the teacher's guide was essential; she could not have done without it. It provided a lot of background knowledge and was laid out nicely. The materials on sands were most helpful to her students, and they learned much about sand, beach, etc. The best and most effective lesson was the sand part, and the least useful and effective lesson was the organisms on the beach lesson (the topic was not as interesting to them).

She found the magazine assessments were really good for assessing students' content knowledge, and especially liked *Sandy's Journey.* That was really useful for a literacy assessment, testing both content knowledge and students' literacy development. The assessments did help the teacher to see a broader picture instead of just a few who raised their hands all the time. Since the teacher had used a GEMS Unit before and she used inquiry-based skills in her teaching, she found it easy to use the Unit. She thought this might be difficult for someone who was used to teaching chapter by chapter.

In her opinion, the Unit was well balanced in giving attention to both science and literacy, and it was very effective in helping students learn the concepts and skills related to science. The teacher would use the Unit as a main curriculum again. The teacher would definitely use the magazine assessments in the future, since the students loved them and did not even realize that they were being assessed.

Problems in implementation. The Unit underestimated the time needed. Materials that were supposed to be covered in 3 hours actually took weeks to cover.

One of the writing assignments (where they write their own glossary) was too hard, and not all students could finish it. Students could not get the erosion concept even after she covered it extensively. They also needed extra support in using the materials on hand lenses, water, and the water dropper.

Student learning. *Shoreline Science* was excellent in helping students learn. The teacher thinks the students learned the most in science content and in writing, though not much in reading. The Unit taught students new skills in science—measuring, observing, making notes, and making predictions—and students learned

the first two skills very well and found the last two challenging. Students learned how to write a report and how to write a glossary.

Most effective for. The Unit was most effective for her higher-achieving third-grade students.

Factors for successful implementation. She contributed the success to her previous experience with the GEMS Units and being familiar with the inquiry-based skills.

Recommendations and suggestions.

- leave out the activity of making a big class book out of all the individual reports, because it was just an added step that took another week while it did not really change students' knowledge level
- develop concepts like erosion a little bit more as a vocabulary term
- have a stronger emphasis on literacy and more development on reading
- provide videotapes on the experiments (how do to them) for first time users
- have some kind of accommodation for special learners. There were enrichments, but there were no suggestions on how to make it easier
- break the materials from the big binder (teacher's guide) into segments for easier use
- bind readers differently so they do not fall apart
- have bigger readers so it is easier to share books
- provide more books so students can do more independent reading

Other comments on instructional changes and teacher learning. The teacher liked some of the cooperative learning practices, such as the Line Up. She thought it was a good way of getting information from her students, and it was fun. She would also adopt the practice of having students write what they had learned every day on sentence strips so they could rephrase and reflect on what they had learned, besides teaching her students how to write a report. She gained a lot of content knowledge from using *Shoreline Science* (e.g., about sand).

Profile for Teacher 7

Classroom background. This teacher used the *Shoreline Science* Unit with her third-grade students. Five percent of her 20 students were estimated to be below grade level in science and 95% at or above grade level; for literacy knowledge, 80% were below grade level and 20% were at or above grade level. There was one ELL student in her class who was proficient but did go to a reading specialist. She also had 2 special needs students who were below grade level. They were mainstreamed and sometimes came in with an aide.

Teacher background. This teacher spent about 100 hours in science professional development in the past 12 months and about 300 hours in the past 3 years; and she spent 60 hours of literacy professional development in the past 12 months and about 300 hours in the past 3 years. She has an MA in English Literature & Education. She holds a California Multi Subject certificate and a New Jersey Elementary Lifetime certificate. She ranked herself as "high" in her computer experience, "moderate" as a science teacher, "low" in Earth science, and "high" as a literacy teacher.

Unit implementation. The teacher used the *Shoreline Science* materials three to four times a week for about an hour each time with some exceptions that ran longer. She started the Unit in September and was almost done on December 15, 2004. She found the materials to be creative, visual, and useful. The teacher administered both the pre- and posttests to her students, and also used the *Beach Bucket News,* embedded tasks, and critical junctures for student assessments, besides some teacher-designed worksheets for note-taking purposes. However, she used her district's writing rubric for scoring students' writing after making sure that the contents were covered. Assessments were used to guide her teaching and find out who needed extra help since she did not need to give a science grade to her students.

The teacher implemented everything from the *Shoreline Science* Unit except that she changed the instruction on the postcard activity. In place of writing a five-paragraph postcard, her students did a shorter postcard plus a follow up writing on the storyboard. She used the ELL considerations for her lower performing students and sometimes for the whole class since these considerations were a little bit more creative and visual. She also did not do the making of the book with folding paper. Besides using the *Shoreline Science* materials, the teacher supplemented the Unit with some materials on "schoolyard ecology" before using the *Shoreline Science* materials,

as for her, the third-grade science standard is habitats and adaptation, not rocks and erosion. The teacher used all the reading approaches in her class, depending on the book she needed to use.

Student engagement and motivation. The students were excited about doing hands-on science and they talked about it all the time. They were very engaged and wanted to learn and asked, "What are we doing in science today? When are we doing science?" The students especially loved the readers. They would look at them even when we were not doing science.

Students wrote a lot about science in their "reflection of the week" activity. They also wanted to look for rocks (topic erosion) at home to try to rub them together. They started to talk like scientists. One student was so motivated that he wrote a page and a half about density, water tension, and surface tension.

Unit quality, usability, and utility. The teacher thought the quality of *Shoreline Science* was fantastic, especially the readers. The materials were right on grade level, and the activities were equally appropriate for all her students. The instructions were clear for the students, and they understood them well. The teacher liked the fact that the Unit included so much writing. She wished there was a series of readers like *Shoreline Science* for her students on every standard in science. She also got positive feedback from the parents.

She identified the pretest as an eye opener since her low-performing students could not do it. Her favorite test was *Joan Visits the Beach*. The sand journals were the most useful activity because it really offered the students a chance to read a book, take some notes, apply their knowledge, do some investigation, and then wrap everything up with a great hands-on journal. She found the activity "cutting up the globe" to be least useful, though she liked the idea, it was difficult to implement. The teacher liked the graphic organizer, the note taking sheets, and the science journals, as they were very hands on. Students' favorites were *Gary's Sand Journal* and *Sandy's Journey to the Sea*.

The teacher thought the Unit was well-balanced in giving attention to both science and literacy. She would definitely use the Unit again because of the excitement level she saw in her students. She loved the activities, and the Unit gave students a lot of writing practice, as well as some great material to read. She would use the Unit as a main program.

Problems in implementation. The teacher was pressed for time, for both time to prepare and time to teach. For an example, she did not think enough time was allocated for the process of student writing, teacher reviewing, and student re-writing.

The students did not like the literacy test. It was challenging and sometimes too technical for them, especially the part where they circled words and the closed activities.

Student learning. The teacher noticed that her students started to use scientific terms in both their conversation and writing. Student knowledge grew from "had no idea why things were the way they are" in the pretest to being able to talk about adaptation, etc., in the posttest. Students learned how to take notes, how to ask a question, how to find answers, how to organize their writing, etc. They started to try to apply what they learned to other things and come up with a reasonable answer, though it was challenging for most of the students at this grade. The teacher noticed an improvement in their understanding in social studies and exposition.

Most effective for. The Unit was most effective for her higher performing third-grade students even though all students were engaged.

Factors for successful implementation. She thought the readers led to the successful implementation of the Unit in her classroom, since students had something concrete with real pictures to look at.

Recommendations and suggestions.

- include more pictures in the teacher's guide
- need more space for writing in the student books/worksheets
- prefer to have the Unit aligned with the standards she had to cover
- need more time for the Unit
- the literacy test could be shorter
- some of the writing activities can be combined

- have one copy of the materials on organisms that is laminated so the teacher could have different groups pick different organisms to work with instead of all groups picking the same ones

- more books on different topics (standards) within science would be nice

Other comments on instructional changes and teacher learning. The teacher started to use the storyboard and the postcard idea in her regular lessons because of the *Shoreline Science* Unit.

Appendix D
Shoreline Science Coding Summary

This appendix presents a summary of teacher responses during the interview to the research questions. The structure of the summary mirrors the individual profiles.

Background on Classrooms and Teachers

Three teachers taught 2nd grade, and four teachers taught 3rd grade.

Students' level of literacy proficiency.

- Above: Two teachers reported to have 10%-25% of their students above grade level.
- At: Two teachers reported to have between 25%-30% of their students at grade level, four teachers reported to have between 59%-80% of their students at grade level.
- Below: Four teachers reported to have between 10%-41% of their students below grade level, three teachers reported to have between 50%–100% of their students below grade level.

Students' level of science proficiency.

- Above: One teacher reported to have 20% of their students to be above grade level.
- At: Three teachers reported to have between 1%-15% of their students at grade level, three reported to have between 70%-90% of their students at grade level.
- Below: Three teachers reported to have between 25-30% of their students below grade level, three teachers reported between 90%-100% of their students below grade level.

Hours spent on science professional development. The amount of hours spent on science professional development in the last 12 months and past three years:

- Six teachers spent from 0 to 40 hours in the last 12 months. One teacher spent at least 100 hours in the last 12 months.

- Two teachers only spent 10 to 15 hours in the last three years, the other 5 spent between 100 to 300 hours in the past three years.

Hours spent on Literacy professional development. Amount of hours spent on literacy professional development in the last 12 months and past three years:

- All teachers spent between 0 and 30 hours in the last 12 months.
- One teacher spent only 7 hours, but the 6 others spent between 60 and 300 hours in the past three years.

Highest level of education. Four teachers had a bachelor's degree. The other three had a master's degree, of which, one had 40 hours beyond that.

Teaching credentials.

- Teacher One: Cross-cultural, Language and Academic Development credential (CLAD), California Preliminary (5 years)
- Teacher Two: Bachelors, Master of Education (Illinois)
- Teacher Three: Elementary self contained, gifted and talented (Texas)
- Teacher Four: Bilingual Cross-cultural, Language and Academic Development credential (BCLAD), California multiple subjects for K-5, supplemental authorization to teach Spanish through 9th or 10th grade, National Certification
- Teacher Five: Certification in K-6 (common branches in New York State), certification in New York State
- Teacher Six: K-8 standard teaching certificate, provisional gifted endorsement (Arizona)
- Teacher Seven: California Multi Subject, New Jersey Elementary Lifetime, National Board Fellow

Unit Implementation

Days of using the Unit in the class.

- Three teachers mentioned they used the Shoreline Science Unit 3 or 4 time s a week. One teacher mentioned that she first used it 3 or 4 times a week initially, but that she used it everyday near the end of the unit.

- Two teachers mentioned they used the unit everyday. One teacher mentioned that it varied a lot, from teaching it not at all in a week to teaching it everyday.

Materials used. In general most teachers indicated they used everything.

- Readers: All teachers used the readers.

- Teacher's guide: All teachers used the teacher's guide. One teacher specifically stated the following:

"I used the teachers' guide, the pictures helped me. The guide was imperative. I could not have done it without the guide. And the pictures actually helped me do a better job, because it gave me an idea of—when it said, 'Have the kids draw and put the concepts on the wall,' that really made a difference for me to see, 'This is what other people are doing, this is what I could do.' And sometimes I would read it, and because it was new curriculum, you miss a lot of things the first read. So I know I read them more than once, each of the descriptions. When I looked at the pictures, it helped me understand, because sometimes I'm pretty visual. It made a difference."

- Other: Two teachers mentioned specifically using the worksheets/handouts.

Activities used.

- All the teachers used the assessments: embedded assessments, magazine assessments and pre- and post-assessments were mentioned. When asked about the critical junctures, a lot of teachers needed a little description to remember what they were. Most teachers said they used them.

- Five teachers used the ELL considerations. One teacher used it for her 2nd graders even though they were not specifically ELL learners. Another teacher did the same thing; she used the ELL considerations for her slower students and ESL Students. A third teacher used the ELL considerations for her slow students since they were a little more creative and visual.

One teacher only used a few since she only had Level 3 ELL students in her class; they did not need that much attention. Interesting was the fact that one teacher remarked that she used the ELL considerations but that she did not use them exactly. This influenced how she taught.

- All teachers used the home activities: Not everyone used all the home activities though. One teacher did not do all the home activities because she was out for three days. Another teacher ran out of time and did just one home work activity. A different teacher only used two of the home activities since she had to translate them into Spanish (for the parents) and this took a lot of work. In fact one teacher said she made up her own home activities. The following is an example:

> "I made up some of my own home activities. When we were looking at sand grains and we put them on the cards, sand on the cards, what I did was I had the kids take home a hand lens and a baggie, and I made up a little handout for them and I gave them some file cards, and I said, 'Find things at home that you think might have grains.' And we defined grain as a small tiny bit of something. 'And make your own card, and describe it.' And we talked about properties. I did that once. We wrote about visiting a beach once, and they drew a picture. And I put together something else that I gave them to take home."

Reading approach used. Two teachers indicated they used all three reading approaches: independent, paired, and shared. Some mentioned they used some independent reading. Most teachers said they used mostly paired and shared reading.

- Independent reading: Four teachers mentioned they used some independent reading. For some of the writing projects, one of these four teachers let the students read individually. Another of the four teachers said she used independent reading with only the higher students. The third of the four teachers remarked that she let her students read independently during the independent reading workshop that they had. Two teachers mentioned the following reasons they did not use independent reading: there were not enough books and most students could not read the materials independently.

- Paired reading: Five teachers indicated they used paired reading.

> "When we did the Readers, this is how I always did them. I'd pair the children up. Then I would have one book. We would preview and predict and all of those things. Then I would read it to them while they read along with me. We would talk about vocabulary. We would try to tie it all in. Then the next day, we would read it again, but they would

take turns reading parts of it… I pretty much always did it that way. That was back in September and the reading was pretty difficult for them."

- Shared reading: Six teachers indicated they used shared reading. Following are some of the reasons:

"We did it with the whole group because the reading had so much vocabulary in it. When I tried to have them read it by themselves, they would all go read it, and they'd read the words, and come back not knowing what they'd read. So we needed more. That way when we were all together they could ask questions. Usually, it would be a question somebody else had as well."

"I have 25 students and they gave me 10 of the readers. To have them read independently was just not feasible. The buddy reading did not always work because of the pairing. I have a lot of chiefs and not too many Indians."

- Other: One teacher read the reader first to her students before the students read them. This worked better with the 2nd graders. Another teacher said that it depended on the book as to what reading approach (paired or shared) she used:

"I think depending on the book that we did. Some of the stories lent themselves to reading and discussing as a whole class. Some of the reading I made up an organizer for them and as they read they had to take notes. They did that in pairs. We have 20 students in the class. A couple of times a week I have some parent volunteers and we broke up into small groups. So I might read with six of them. When they came to my station that was what we were working on. We'd work on that."

Use of assessments.

- One teacher used the assessments as instructed in the unit. However, she did not grade the embedded tasks and magazine assessments with the existing rubrics (4 point scale). According to her, the grading system and the ways to grade them were very confusing and hard (graded in different areas). Additionally, it was on a different scale. The teacher used a 5 point scale. She did grade the pre- and posttest as explained in the *Shoreline Science* unit.

- Two teachers stated that they used the assessments for giving grades. One of them used the assessments for a specific language and reading grade. She made her students make a class rubric together. See the following example:

"I recorded grades in my grade book based on those Rubrics and what they needed. Of course, we talked about it extensively before. We posted a class Rubric. It was an abbreviated Rubric of what you needed to have in your book if you wanted to get the highest score….I think if children generate the Rubric they buy into it more. If I say, 'If you want to get a six, you have to do that. If I have a scale from one to six, what do you need to have in order to get a six?' So they'll tell me, and I'll post that on the chart. You to have a table of contents and you've got to have the pages in order. We represented it on their class Rubric."

- Two teachers used the assessments to get an idea of who really understood it and who did not. One of them mentioned that she had not used the post assessment yet, but as soon as she did, she would look at that one to see which of her students did understand it.

- One teacher used the pretest to guide her instruction. See the following for her statement:

"The pretest was just as it's stated. In terms of using it to drive the instruction, I did actually grade those. I did not grade the post test, because I think I had to have them in. They took a long time to score. Basically, it told me that they had little to no prior knowledge, so that did lead me in terms of the lessons. Sometimes you can kind of skip over some things or others, and I just knew that they did not have the prior knowledge and that I needed to cover things thoroughly. I also knew that they really lacked a lot of the vocabulary, the reading ability, that those things I really needed to focus on."

- A second teacher said that she used the assessments to guide her teaching (also because they do not have a science grade). If she noticed that students couldn't answer the questions after the instruction, the teacher would adjust her way of explaining the contents, thinking she was not presenting the contents correctly.

Changes to the Shoreline Science Unit.

- One teacher remarked that she changed some things in the *Shoreline Science* unit. For example, she changed the partner sharing portion. She did those differently because her students had other ways of doing them already, and it would take less time to do things they were already familiar with. She had the students do these activities individually. Furthermore, there were some other minor things that she changed.

Student Engagement and Motivation

Teachers said that their students took an active part throughout the whole unit. The following are examples:

> "They really liked the Jolly Rancher one. That was really cool. They really got that. In fact, they'll even come to me now sometimes with chalk. 'Look, we made sand.' It's all the chalk dust, so they really got that. They really enjoyed the globe activities back in the beginning where they tried to go around the globe with their one finger over water, and then over land. They remember that really well. The oil spill. They were real into that."

> "Where I said before that they have a hard time because they take things for what you say, but at the same time they'd go out to recess, and they'd come in with a rock or something, and they'd say, 'Look what color I see. It must be this or that' Whereas before they'd see a rock and think, 'This is pretty.' So it's like I know that they were really grasping and taking it beyond the classroom….A lot of the students brought in items from home that would have to do with things."

> "I'll tell you what else, we went on a field trip to Second Nature Center here, we have a big city park here in Long Beach and there's a little section with a nature center. It was so interesting, because there were a lot of the same vocabulary words like 'adaptation,' and the kids were really engaged. Even the tour guide was like, 'Oh my gosh, your class is one of the best classes I've had,' and I was kind of like, 'Wow,' because they're so low academically speaking, but since they had all this input and such great activities—not all of them, but the higher ones really learned a lot and are starting to apply things in different areas. So that's good. I just want to say that it's been really great, and it keeps coming up, the things we learned as they're recycled and becoming more embedded. That's probably it."

Students' reaction. All students loved the Shoreline Science unit. Following are examples of students' reaction to the unit.

> "They would talk about what they like. In January we had a writing test. This had nothing to do with science, but they had to write and tell about school. Almost all of the students' favorite thing was science."

> "They loved it. They really enjoyed it. There was not one bored person in my room. I mean that literally. It really came alive for them."

> "It was really high interest for the students, really motivated, at least my group. They were very engaged. They were so excited to be doing hands-on lessons and things like that, they love."

"My students absolutely loved it, and you would not have believed it because at the very beginning—I had to start it right off, like the second week of school. After the first few lessons, we were in the middle of the second or third hands-on lesson, and some kid was like, 'Wow, this is the best class ever, we never got to do anything like this before!' They were always making comments like that. It was very interesting. I have an amazing group this year, in as much as they're really low, it's just a very special group. They're pretty much engaged in a lot of things. And I brought that to it, it would be like, 'Today, guess what we're doing in science. We're going to learn about an oil spill!' And they would cheer and just be so into pretty much everything. So for my students it was very engaging."

"They liked doing the journals, but some of them got so into looking at the rock boxes—it was like they had not seen anything like that before."

"Yes. They just could not wait. Every day, 'What are we doing in science today'? I think that they really just could not wait. Every day they were ready to tackle it and could not wait to see what we were doing next because there was stuff."

"The best thing about that program was the readers. They were phenomenal. My kids loved them. They would look at them, even when we were not doing science."

"I do something called 'looking back at the week' with my students. They do a reflection on Friday of everything they've done during the week. They have to write about one thing in particular after they go through. They can't just say, 'We took a math test.' I say, 'Your parents already know you took tests. What did you do in class?' So for that entire time most of the students wrote their paragraph about what they had done in science. I think my kids were just so excited to be doing a hands-on science because they have not done it. They were just talking about it all the time. One of the things that we had to do in class was to use a pumice stone to rub together to show erosion. We talked about how that's not how it is in nature. You're not just going to rub two rocks and all of a sudden you have all this. But the kids said, 'What if there is a rock outside that we could do that to? Could we try it?' So I said, 'Sure. If you bring a couple of rocks in from recess.' So of course every child in class came with every pocket filled with rocks and they were all… we had been in the science center. One mom said her son came home from school and spent all afternoon out in the backyard rubbing rocks together. Somebody brought in a Ziploc bag, which I think was dirt, but… 'This is the sand I made.' I'm sure it was the dirt that came off the rock, but they were just bent on trying to prove that that could happen. The language was great, learning how to talk like a scientist. I started to see that in their writing, using the words observe and investigate."

"One of my students, when we did the oil investigation homework investigation, came back with a… As I mentioned they only had to write a paragraph, in third grade pretty

much six sentences is acceptable if they have detail. He came back with a page and half written using the word… he's a very bright student, but his parents are not the kind of people who would write it for him. His mom is an engineer. She sat down with him and they talked about it. He talked about density and water tension and surface tension. He was reading his essay to the class and half the class was looking at him like, 'What are you talking about?' He just was so excited. He was like a sponge. He was so excited to be learning those words. He did a rough draft. He did a four square organizer. It was not as if, his mom was telling me, she said, 'You need to sit down and do it this way.' He was just totally motivated on it. It motivated students is what I have to say. It really motivated them to do a little more than they usually do…more of them."

Unit Quality, Usability, and Utility

All seven teachers found the curriculum/teacher's guide useful. It was clearly written and planned out, gave lots of ideas, strategies, and background knowledge, and laid out all the experiments.

The timeline was helpful because it made teachers stay on task. Someone else mentioned that she liked how the reader was divided into before reading activities, during reading activities, and post reading activities.

Readers. All teachers found the readers very useful.

- One teacher said that she gave some of her readers to her other second grade teachers that wanted to use the organism reports.

- Another teacher commented as follows:

"The readers were great because there was a lot of information in them. They were really well organized, the pictures were great, the vocabulary—for some kids it was challenging, but there were definitions. We could get the definitions. The idea of using and building a glossary with the concepts and the terms that were used was really terrific. The whole idea of evidence, and what evidence means and how to use it in those books, and how to use it when we made the web and looked at the models. It makes the word 'evidence' come to life for these kids… I can now make a connection all the time too. 'Remember when we looked, when we found evidence of what the sand was composed of'? So I link back to that all the time, so they have a very concrete example of what evidence is …we do workshop model reading, so the kids are picking independent books based on interest. We do not have a textbook or an anthology or basal reader; the kids just choose what they want to read…it was one of the best ways for me to begin reading at the beginning of the year. Especially with a group of kids that are strugglers and are really challenged. And it would only be better with a group of kids whose skill levels were much higher."

- Again another teacher mentioned that she liked the otter book the best. Only one of the books was a little too hard, but it was used as a resource book only, so that was fine.

- One teacher was so enthusiastic that when she put together her next unit, she specifically went out to find books at their grade level as the GEMS unit had done and tried to find at least 10 of them so she could have her students work in reading pairs since she found that very beneficial.

Reader reading level.

- One teacher said that she thought that the reading level was exactly right for her students.

> "It was nice to have the multiple copies for the paired reading, especially with my group that has varying reading levels. They could be paired with that. I really felt they were right to the point, really good illustrations, photos, depending on the book. Most of them, for English learners, very good. Really looking at the text, they did not have a lot of compound sentences. Pretty much simple sentences throughout, so that really helps the readability level."

- Another teacher mentioned that the reader was too hard for her below basic students.

- It was challenging even for the high readers for a different teacher. She went through the books twice and referred to the books during the activities. This made her students internalize the terms and the concepts. The lower readers could still not read those, but they were using the terminology in reports and on posters and could understand those.

Reader science level.

- One teacher mentioned that the readers were very useful and more manageable than the science text.

Unit Balance. All teachers thought the unit gave well-balanced attention to science and literacy. Only one teacher mentioned that there could have been a few more literacy type activities in the unit such as fluency training or taking big words apart.

Usefulness of assessments to measure progress. Three teachers thought that the assessments measured more growth than progress.

- Only one teacher said that the pre- and posttest did not really measure progress, they measured growth. She thought that the magazines did measure progress though.

- Another teacher remarked that the assessments did not exactly measure progress since they were not assessing the same thing each time. Different topics were addressed in different assessments.

- A third teacher said that Lesson 2 was not that great in terms of assessments. It did not really measure progress, it gave her an idea that her students could organize information.

All teachers stated that the assessments measured growth. Two teachers mentioned the pre- and posttest in particular.

- One teacher stated that either the students understood the concepts and that they could represent them properly on the various instruments or they could not.

- Another teacher remarked that it gave her an idea of how much the students understood about what was taught and it gave her an idea of how well they could express themselves in writing and if they were using some of the vocabulary of the unit. For the pre- and posttest she did not know yet since she had not finished. But she thought that it would give her an indication of how much more background knowledge they had acquired in science and how much their reading had improved.

- A third teacher indicated that she saw from the mini-books which students were able to understand it and who were not.

- A fourth teacher mentioned the following:

"A lot of those assessments. They were not exactly assessments, but they were ways for me to say, 'Can this student explain something? Can they explain something verbally to each other?' I think it's great. The first two."

- Another teacher realized through the assessments that some students did not understand the topic. They helped the teacher to see a broader picture instead of just a few who raised their hands all the time.

- Another teacher said that it guided her teaching and that it showed her who needed more.

> "Even though I did not have to use them as a grade they helped me, as I said, guide my teaching and it let me see who needed more. There were kids that I found other information for because they were so involved in what they were doing. Who came in with something they got off the Web."

Usefulness of assessments to measure progress for EL's. Four teachers said they thought the assessments were useful to measure EL's progress.

- Only one teacher said that the pre- and posttest did not really measure progress, it measured growth. She thought that the magazines did measure progress though.

- Another teacher mentioned that they were very useful and that the rubrics were helpful in scoring them.

- A third teacher stated that as long as you give the EL's individual support, they do well on the assessments.

- A fourth teacher indicated that she saw from the mini-books which students were able to understand it and who were not. Even the extremely low students seemed to gain some knowledge.

Using the Unit again? All teachers would use the *Shoreline Science* Unit again if they had the choice. Some reasons for this were:

- The kids loved it, they learned a lot. It was effective and the students were very excited.

- It was outstanding in the way that it sequentially presented science concepts and discovery activities (inquiry-based learning).

- It not only provided facts, but it built the concepts up.

- Because of the mini-books, writing was such a great assessment, it was authentic and individual. There was lots of writing practice and great material to read.

Four teachers would use it as their main program. Three teachers would use it as a supplement for the following reasons:

- not being allowed to use it as a main program (2 teachers)

- not having enough literacy in it (1 teacher)

Parts of the unit that would be most popular to use again were:

- report about organisms
- concepts from all the lessons, but not going into them in depth
- use the lessons that really address the standards
- create a knowledge map
- part of the Jolly Rancher section
- the mini-books

Units liked by students:

- make sand samples, exploring the sand, the *Shoreline Science*, paragraphs about sand, Sandy's journey, sending postcards
- organism report
- science journals
- science tools: magnifying glass, using charts and graphs
- study about organisms
- Jolly Rancher
- globe activities
- the oil spill

Units disliked by students:

- vocabulary part (science words and everyday words), especially when the teacher wrote down the definition
- discourse circles
- copying and writing of the report
- the literacy test

Problems During Implementation

- The time estimate was unrealistic, too much contents for too little time.
- The first two units could be expanded a bit and the other two could be shortened or eliminated. In general the unit took too long.
- The report writing took much longer than the estimated 3 hours, it took weeks. Instead of making a big class book of all the reports, maybe just do the individual reports.
- One teacher indicated that her students just had 2nd grade problems with listening that had nothing to do with the curriculum.
- Another teacher said that the higher level analyses were really hard.
- Someone else mentioned that the students had problems with writing the paragraph. Not all students could finish, it was a stretch and really hard for them.
- One teacher stated that more background information on some things would be helpful. For example, if it said to find more information in a certain section, it did not say where that information was to be found.
- Another teacher said that it was a lot of reading and that it would take a couple of readings to make sure you knew all the steps and had all the materials ready.

Four teachers stated that the students needed some support. Following are the topics mentioned:

- reading and writing
- hand lenses and water droppers

Student Learning

Learning reading concepts and skills well. Most, but not all teachers said that the students learned the literacy concepts and skills well.

- The students learned how to use the index, the glossary, and the way to look through a non-fiction book. They also learned how to write a report, with topic sentences and how to stay on topic.

- Another teacher mentioned that her students learned how to make inferences and write little research reports.

- One teacher noticed improvement in social studies and expository writing.

- One teacher said that re-reading, looking for evidence, proving ideas, and answering a question at the end of a paragraph really helped her students. She liked the way it was done with the different colors; the idea of linking concepts with a color helped organizing students' thoughts.

- The way the students learned the features of non-fiction text and the writing of the report with looking through the index and table of contents really helped the students.

Learning reading concepts and skills that did not work well.

- One teacher mentioned that she did not like the fluency, she would rather use Houghton Mifflin for that.

- Another teacher stated that she did not think the unit was strong in literacy. The writing (report and glossary) was stronger than the reading.

- Another teacher said that using the *Shoreline Science* Organism book as a reference book did not work that well. The students did not really go to the paragraphs to read, they just went to the little information boxes and took the things that were easy to find.

Learning science concepts and skills well. None of the teachers said that their students did not learn the science concepts and skills well.

- It was remarked that the way the curriculum was built really helped the students learn the science concepts (even though it was hard for some).

- One teacher mentioned that her grade level and above grade level students really learned the concepts well. "They can teach this unit," she said.

- It was a little bit difficult for some students because they did not have any prior knowledge, but they connected the things they learned to things they saw on the news and wherever they went; they seemed to remember. This was mentioned by at least two teachers.

- Another teacher mentioned that her students can define a lot of concepts now.

- Someone else thought that the sand unit was the most effective.

Student challenging reading.

- One teacher indicated that the students were not really challenged by the reading. The literacy part was just right on.

- Three other teachers said that their students were challenged by the reading in the *Shoreline Science* unit.

- One teacher knew her students were challenged because of how much they learned and by how much their thinking had expanded.

- Another teacher mentioned that the organisms book was a little technical; that was a little challenging for her students.

Student challenging science.

- One teacher said that her students were not really being challenged by science since there were so many things they did not know.

- Four teachers reported that their students were challenged by the science.

- One teacher remarked that the students had to think in a way they had not thought before. Some of her students were developmentally "not quite there yet." On an individual basis, she said, having to think in a different way is probably the biggest challenge.

- Someone else stated that science is new in her district and that it is only the second year of using a science program. Therefore the processes of science were hard for her students. Measuring and observing were new skills, and writing down observations and making predictions was a challenging job.

- Another teacher said that her students did not have a lot of worldly knowledge and that they do not investigate often. Asking them to apply what they learned and think about it and to come up with a reasonable answer is challenging for most of the students in her grade.

Unit Most Effective for

For 2nd graders, the Unit was most effective for those who were at or above grade level. For 3rd graders, the Unit was effective for all students, especially the high-achievers and those at and above intermediate language level.

Successful Factors

Reasons why the implementation of the unit was successful:

- engagement of the students
- mix of activities and the way the sequence of the lessons was designed

"In other words, we have an introduction. We have the book, and then we have some hands on activities. Then we have reflections about the activities. Then we have the applications of the concepts to real life. I just think it's been created so carefully."

- group activities versus partner activities
- creating a structure by using concept maps, diagrams, etc. for students to internalize concepts; content area information is needed
- the integration of literacy and science
- teacher enthusiasm and willingness to do the preparation work
- building not only background knowledge, but also building concepts
- going from concrete book ideas toward real experience
- familiarity with the GEM's unit and format
- the readers were concrete with real pictures to look at

Recommendations and Suggestions

The following suggestions were made to change the unit/program to better meet all students' needs:

- need more ideas for the teachers on how to modify assignments for lower level students and special learners, especially regarding reading

- highlight the must-dos or prioritize lessons to guide teachers with time constraints, and to guide teachers who use it as a main or supplement

- have books of a lower reading level, for example: more pictures cards for support, more streamlined lessons that are not that long and not that many; a few different vocabulary words each time instead of all at once

- include more pictures for visual learners

- need more discussion to explain concepts to 2nd grade students

Suggestions on content changes.

- need more activities for the lower students

- need more attention to vocabulary development

- shorten the discussion lessons or include other activities so that every learner is engaged; shorten the group writing exercises

- eliminate the activity in the class report where the students had to just copy the paragraphs

- include more writing on a day-to-day basis (e.g.: reflection journals)

- consolidate some of the writing

- have more lessons in writing: writing should be more of a direct component of the lesson on an on-going basis and there should be more lessons relating to the vocabulary since it was so content specific

More suggestions were as follows:

"Include a vocabulary activity or include some vocabulary words or include some reading practice activities where the language or the concepts are used. Maybe have a cause and effect comprehension worksheet or a problem and solution worksheet that relates to whatever concepts we're teaching. Like we had the lady visiting the beach on the test. You could expand that and have reading comprehension activities that relate to what you presented."

- change Activity 4 for a more exciting activity, plus more information for the teacher to understand Activity 4

- implement the rock boxes separately because the students had never seen it before and were amazed by it

Improvement for materials.

- have a reader for every two students
- bigger books to be used for shared reading
- bind the readers differently so they do not fall apart so fast

Changes for the teacher's guide. Almost all of the teachers mentioned that they would like to make changes to the format and layout of the teacher's guide. Following are some of the comments mentioned:

- Different lay-out for the teacher's guide. Now it was a lot of flipping back and forth. Nicer to have the directions and the experiment right in the lesson.
- Some of the information in the teacher's guide did not always connect. Information found in the overview was different from the written information later on.
- There were sections that referred to the teacher information that did not exist.
- The teacher's guide was hard to use, not user-friendly. It was easy to get lost. Maybe having it in segments instead of one big binder would help.
- Include more pictures.
- Include more anecdotal information and background information.

Teacher instructional changes as a result of the Shoreline Science Unit. All teachers stated that they had implemented changes to their regular practices as a result of the *Shoreline Science* Unit.

- One teacher will be more thoughtful about writing the science vocabulary on the board separate from the everyday ones. This teacher will also let her students get more involved in doing the science experiments before the lesson instead of after the lesson.
- Another teacher liked the glossary on the wall, liked using the science terms, and defining them more clearly with everyday terms and

examples. She will use the shared partner activities where the person shares and refers to the name of their partner. She will use concept maps since that helped the children's understanding of making a concept map and using the concept map for writing. It was a springboard for instruction in other topics. She will also use the report writing, using the same format for creating a Four Square Organizer and plugging information into the various topics about the subject or subtopics.

- Another teacher liked the way the discourse circles, and the think/pair/share, activity routine were organized. She had done these in the past. She also liked the use of a web when showing how sand was created and the composition of sand which was very visual and meaningful to the students. She also liked the routines that were introduced.

- Cooperative learning practices is what another teacher will adopt. The Line Up for example, where the students showed a picture and they had someone stand in front of them and ask about their picture, then the line moves. She thought that was a good way of getting information out and more fun. She will use it now with other things.

- One teacher is going to do more of the storyboard. That is how she wraps up the unit. The students think like scientists when they brainstorm as a class and go through all the different things they did and they complete this with an ending storyboard. She might also use the postcard idea since it had a lot of writing and that was very good.

- One teacher said that she probably has infused some things from the unit. She also incorporated some of the literacy parts, like the writing component.

- Implementation of the concept board (wall) was mentioned by one teacher. For example:

"In our reading series we also have what's called a concept question board, but really as part of... I'll always do the closure to the lessons and have kids orally share out what they've learned. But in the program, after each lesson, they came up with key ideas—you started together, more teacher-led, with concepts to summarize what they learned in the lesson. Every time you wrote it down on a sentence strip, and then you'd have a student or a couple of students illustrate that concept. So I've started doing that with different lessons, with the science or social studies I've continued on with."

Appendix E
Terrarium Investigations Teacher Profiles

Profile for Teacher 1

Classroom background. This teacher used the Terrarium Investigations Unit with his 2nd grade students. Out of the 15 students, 30% of them were estimated to be above grade level, 50% at grade level, and 20% below grade level in science proficiency. For literacy proficiency, the estimates were 40% above, 35% at, and 25% below grade level. There were no ELL or special education students in the class. The below grade level students got some extra help from the teacher and teacher assistants.

Teacher background. This teacher spent 24 hours in the past 12 months and 100 hours in the past three years in science professional development; he attended 20 hours of literary professional development in the past 12 months and 60 hours in the past three years. He had a Bachelor of Science in elementary education and a K9 teaching certificate from Illinois. The teacher ranked himself as "high" in his computer experience/knowledge, as a science teacher, in earth science knowledge, and in expertise as a literacy teacher.

Unit implementation. The unit took the teacher 43 or 44 days to complete, and he used the Terrarium materials almost every day for an hour. The teacher used almost everything that was sent to him, except that he was not sure whether he gave the pretest. He used a few things from the ELL considerations, e.g. the specimens, a few home activities. He re-taught some lessons for students who had some misunderstanding or who were absent, by having some small group discussion or mini lessons.

He added in some trade books, some videos like "Magic School Bus," and fit in some activities to enhance the Terrarium lessons or to cover content that was missing from Terrarium, but that was in the district curriculum. Information from the assessments and teacher-student conversation was used to see whether students got the key concepts to guide the instruction and to give students feedback, besides assigning students grades.

The reading approaches used in the classroom differed depending on the type of text and the purpose. The teacher also used "Fluency Oriented Reading

Instruction" which went through four stages—reading out loud, reading along, echo reading, and paired reading.

Student engagement and motivation. Students were far more interested in Terrarium for learning to read, write, and do activities than they were in the other curriculums. They liked the unit and talked about it the whole period. Because of their interest and the relationship between the books and what they were actually doing, they were motivated to want to understand the books. The books helped them to understand what they were doing, as opposed to just reading a book and trying to understand the concepts that a student may or may not ever want to know. Students reacted very well towards Terrarium and parents found them coming home excited and talking about what they were learning in school. Terrarium was very good at holding students' interest in learning the concepts covered in the unit.

Unit quality, usability, and utility. The teacher found Terrarium to be of better quality than any other books/programs he had used before in 25 years. The teacher's guide was very useful, it consisted of lots of information while being well organized. The directions were easy to follow and understand. He liked the shorter readers because they were at the right reading level for the students and because it was easy for students to go back and re-read the readers and take them home. Terrarium gave students lots of opportunities for interaction between the students, the materials, and activities. Students liked hands-on activities and the poster presentations. Some parents commented to the teacher that they were pleased with their children using Terrarium and excited that their children were doing science and literacy at the same time in 2nd grade. Terrarium gave students high or low a lot of development and opportunities to improve and produce quality work.

The teacher found the assessments useful in guiding instruction and in giving student feedback. The scoring rubrics were good. Through the magazine, the posttest, critical junctures, and student presentations, the teacher was very pleased with the students' reading skills and the level of science concepts students learned. He thought his students' writing skills on writing a three- or four-paragraph essay with a topic sentence and three supporting details etc. needed improvement, judged by the school/district curriculum. His students, by using Terrarium, did other kinds of writing, like making charts and labeling and taking data.

The teacher really liked the literacy part, but he thought he would meet only a portion of his literacy goals and a lot of his science goals if he used all three LHS

Units. He felt that he should do a lot of other kinds of reading to supplement the literacy development. In the future, he would use Terrarium as a supplement to an existing program and as a main science program. The teacher felt that Terrarium made this a better teacher and made his students better students.

Problems in implementation. The students had some problems with the decomposition cycle, discourse circles, making careful observations, and basing their conclusions on the observations.

Student learning. Students liked the hands-on activities, and the readers were at the right level. Students improved their skills in observations, reporting, and recording their observations. By the end of the Terrarium Unit, about 75% or more of the students could read the readers independently.

Most effective for. Terrarium was effective for students of all levels.

Factors for successful implementation. The main factors were the combination of literacy development and science, student interest, and having the literacy development preparing for the activities.

Recommendations and suggestions.

- do another check on typos
- need some help for the low achieving student
- have a black line master book and a teacher's manual
- evolution was not in Illinois' state standards

The teacher found both Shoreline and Terrarium of excellent quality. Shoreline was more alien to the students than Terrarium, but students particularly liked sending off letters to family and friends for the sand samples.

Profile for Teacher 2

Classroom background. This teacher used the Terrarium Unit with her 2nd grade students. Out of the 21 students, 25% of them were estimated to be above grade level, 50% at grade level, and 25% below grade level in science proficiency. For literacy proficiency, the estimates were also 25% above, 50% at, and 25% below grade level. There was one special education student in the class part of the time.

Teacher background. This teacher spent zero hours in the past 12 months and six hours in the past three years in science professional development; she attended 3 hours of literary professional development in the past 12 months and 12 hours in the past three years. She had a Master's degree in elementary education and had a National Board Certification. The teacher ranked herself as "high" in her computer experience/knowledge, "high" as a science teacher, a "moderate" in earth science knowledge, and an "expert" as a literacy teacher.

She had never participated in a GEMS workshop before.

Unit implementation. The unit took the teacher nine weeks to complete, with two days off for snow and a few days off for field trips. She used the Terrarium materials every day for one and a half hours to two hours a day. The teacher used everything that was sent to her and followed the teacher's guide closely. She skipped the last home activity because it was in Spring break. For the ELL considerations, she only used them when she thought they were good for all students.

The teacher re-taught a few lessons, e.g. soil (Activity One), that she did not think the students understood after reviewing the information from assessments, and followed with some writing activities. The teacher supplemented Terrarium with some additional materials on worms and isopods, some resource books about centipedes and different animals that live under the soil, and students were also asked to search for facts about animals and organisms while they were in the computer lab. She also gave students some extra assessments that were required by the state curriculum and that would allow students to apply the knowledge in a different way. The teacher used the scoring rubrics provided.

In the classroom, the teacher typically used paired reading almost every time because there were not enough books for everybody to be reading on their own. With pairing a less able student with a more able one, it really worked well. They

would work together and the less able students would get help from the more able students with the vocabulary. Students did independent reading a few times as part of the homework and to share the reading with the parents. And some other times, the teacher would put the strugglers together and read with them in a small group while having the more capable ones reading with a partner.

Student engagement and motivation. The readers were easier for students who were interested in the topics independent of their reading levels with fictions. The students who were usually the least motivated were more motivated by Terrarium because of its scientific contents. Students liked the Unit, they especially liked building the worm bin and building the terrariums. They liked planning how they were going to build their presentation board.

Students were interested in doing hands-on activities and were taking active parts in the activities. The teacher was surprised to find out that they liked doing the notebook and making the glossary pages, and they also performed better than her expectation. Students also did a good job at the discourse circle activities, especially the second discourse circle activity on organisms. They came up with some really good reasons, and almost everybody was able to come to a consensus; even the stubborn students agreed.

Unit quality, usability, and utility. The teacher found the teacher's guide useful and essential in her implementation. She thought all elements of Terrarium were useful and she especially liked Activity Four the best because students had to use everything they learned in the past. She liked the home activities because they involved parents. She also liked the discourse circle activity because it enabled the students with social issues of getting along with each other to work with the other students.

Terrarium also helped students to apply their knowledge to a different situation and to make connections between classroom activities, e.g. listening to oneself read was similar to listening to your partner, as both required focused attention. While realizing that the more able students had very little difficulty in reading the readers and the less able students struggled, she found the Terrarium readers were easier and more motivating, as compared to fiction, stories for boys, and for students who were interested in the topic, independent of their reading levels with fictions.

Students found it most difficult to understand the lesson on the characteristics of the soil and the easiest to understand decomposition. The teacher would prefer to have more literacy teaching on the content that was covered in the magazine assessments or have the assessments focus on making inferences and decisions to reflect the curriculum. She would also like to see more literacy activities, lessons about context clues, and lessons on new vocabulary. Therefore, the teacher would use Terrarium as a main science program and a supplemental to her literacy materials.

Problems in implementation. The magazine assessment was difficult for students, they knew a lot more than what they put down, they did not seem to understand what the questions were. Maybe it was because they were not used to the format, or because they could not understand the questions in the way they were phrased. Students were frustrated at the beginning because they did not know what the right answer was and because they did not realize that there was not a right answer. They struggled with doing the decomposition cards.

The teacher found the instruction a bit wordy and would prefer to have it shortened while bolding the important points.

Student learning. The teacher found Terrarium to meet students' interest. Students learned an enormous amount of critical thinking skills, over and above the science concepts. Students were well trained in note-taking, making predictions, using diagrams, understanding and using tables, an index, and a glossary. Students learned presentation skills, like using note cards. Terrarium helped students learn all the concepts and skills they were supposed to learn in science.

Students found the vocabulary part in the reading challenging, even though the teacher thought writing activities were challenging to everyone no matter what their levels were. While the lower students wouldn't be able to score as high on the rubrics as the higher students, they had actually stepped a little farther than they normally would. They pushed themselves a little harder because of the nature of the activities.

Most effective for. Terrarium was effective for everybody, while reluctant readers and the boys who usually did not buy into doing things and writing (especially writing about feelings, etc.) achieved above the level than they would have under other circumstances.

Factors for successful implementation. Being hands-on, having students use their thinking skills, and building concepts from the beginning to the end were the main factors.

Recommendations and suggestions.

- Some ELL consideration activities are applicable to all students, especially the one under Activity 1.2
- Include some suggestions on how to support less able students in the teacher's guide
- Shorten the instructions for teachers, make them less wordy and bold the important parts
- Remind teachers to use the vocabulary charts
- Include more literacy lessons

The teacher also did Shoreline and found Terrarium more comprehensive and involved a higher level of thinking, while Shoreline allowed students to be creative and being a creator instead of an observer.

Profile for Teacher 3

Classroom background. This teacher used the Terrarium Unit with her 3rd grade students. Out of her 23 students, 2 of them were estimated to be above grade level, 5 at grade level, and the rest below grade level in science; and 5 students were above, 11 at, and 7 below grade level in literacy proficiency. There were 6 students receiving special education on speech, 3 were emotionally disabled, 5 were learning disabled, and some students attended more than one special education program. The teacher gave them special instruction per their needs during the regular class periods.

Teacher background. This teacher spent probably 15 hours in the past 12 months and 45 hours in the past three years in science professional development; she attended 20 hours of literary professional development in the past 12 months and 30 hours in the past three years. She has a BA degree and will have her MA in reading and literacy June 2005. She is in her seventh year of teaching, and is accredited 2-5 elementary education to teach all subjects. She ranked himself as moderate" in her computer experience/knowledge, in earth science knowledge, in expertise as a science teacher and as a literacy teacher.

She participated in a GEMS workshop before.

Unit implementation. The teacher finished the Unit in the allocated time. She used the Terrarium materials for an hour and five minutes a day, five days a week. She did all the home activities, the magazine assessments, embedded tasks, critical juncture, and both pre- and posttest. She supplemented the Unit with some of the South Carolina Aquarium online curriculum, mainly some adaptation lessons and South Carolina regions activities. The teacher provided extra help for her below grade students. She used the information from assessments to know whether the students needed extra instruction, to give grades, to measure student status and growth, and to give students feedback. In the classroom, the teacher used whatever reading approach that was called for by the Unit.

Student engagement and motivation. The students loved the Unit. Students found the science block to be their favorite Terrarium Investigations of the day. They knew that it was going to be a good mixture of activities and that it would allow them to communicate with their peers, they liked the practice of the think/pair shares and discussion-type activities, and they knew what to expect in those routines. They liked to do the hands-on activities because they enabled them

to compare what they were doing to what real scientists were doing and these activities also motivated them and made it more real for them. Some students commented that the homework was the most fun homework that they had ever received, and they would ask for it throughout the period, "When are we going to have another homework assignment from you?"

Unit quality, usability, and utility. The teacher thought Terrarium was of high quality and was effective in helping students reach the objective and helped students learn. She liked the hands-on activities because they motivated her students and made science real to them. The teacher's guide and the ELL considerations were very useful to the teacher. Additionally, the readers were at the correct reading level yet challenged even her higher level students.

The teacher found the following two activities powerful and useful: where students had to choose and build their own investigation and where students created non-fiction pages in a guided way. These activities really strengthened the students' abilities. The students' favorite was doing things with the terrarium. "The Walk in the Woods" was one of their favorite readers.

The assessments were very useful in helping the teacher to measure student progress and to have a better picture of what various assessments look like. The critical junctions were wonderful, they let the teacher know exactly what the students should have mastered by that Terrarium Investigations and what concepts they should have developed. She found it easy to use the assessments and the scoring rubrics, and she liked the scoring rubrics.

The Unit had a good balance between science and literacy. The teacher would like to the use the Unit again because she found the integration of science and literacy beneficial to her students in fostering their ability to make connections between different content areas and to improve their work quality. The Unit would serve as the main science curriculum, maybe not as the main for literacy as there were required district and state literacy standards.

In this teacher's opinion, LHS' units helped teachers see what integration was all about when it was hard for teachers to see on their own because they taught curriculum in parts, teaching language arts, math, science, and social studies seperately. There was not enough time for teachers to make the connection besides the fact there was no such quality curriculum available. The Unit's integration of science and literacy, hands-on for science, and providing different kinds of

assessments helped the teacher obtain a higher rating as a teacher. These features of the Unit also definitely made a difference in their learning.

Teacher problems in implementation. This teacher did not experience any problems during implementation.

Students problems in implementation. Students had a hard time writing observations even when they were much better at giving oral observations. The teacher thought maybe it was developmental and had nothing to do with the Unit materials.

Student learning. Since the students experienced Shoreline before and knew the procedures, they had no problem understanding directions and accomplishing the tasks as were intended. Every student improved in his/her science concepts and skills. The teacher also observed a very big jump, as big as three grade levels, in students' literacy skills as demonstrated in her State benchmark assessment called STAR. The scaffolding nature of the Unit, combining science, writing, and other literacy components, made it easier for students to make connections to other content areas. The most important learning achievement in literacy was that students improved in both writing and reading.

Most effective for. The unit was effective for all students and all students learned.

Factors for successful implementation. The teacher thought the most important factor was that the Unit gave students opportunities to think on their own and to choose what to write, what went into the terrarium, and what to investigate. The other reason was students' excitement about doing what real life scientists were doing.

Recommendations and suggestions.

- Provide some more examples of what written observations look like and provide more guidance in that area.

- Include more visuals or more examples for teachers to pick what the students need, especially for the struggling students.

- Use different strategies for the before, during, and after the reading of the text activities. The teacher liked the use of the glossary and the components of the non-fiction text.

- Give more support to the teacher with using the rubrics.
- Break activities into small pieces, like doing some collection of data or recording on the chart, do something else, and then come back to it.

Profile for Teacher 4

Classroom background. This teacher used the Terrarium Unit with her 3rd grade students. Out of the 20 students, 20% of them were estimated to be above grade level, 55% at grade level, and 25% below grade level in science proficiency. For literacy proficiency, the estimates were 50% above, 25% at, and 25% below grade level. There were two students who spoke English as a second language, one had above grade level English proficiency and the other was slightly below grade level in academic English. There were about five students who had learning disabilities, three of them were identified officially. The teacher modified her teaching for them and had them get extra help from the teaching assistant and the peers.

Teacher background. This teacher spent zero hours in the past 12 months and in the past three years in science professional development; she attended 12 hours of literary professional development in the past 12 months and 30 hours in the past three years. She had a BA in French and was working on getting a Masters degree. She was endorsed to teach early childhood and English language acquisition. The teacher ranked herself as "high" in her computer experience/knowledge, "moderate" as a science teacher, "novice" in earth science knowledge, and "high" in expertise as a literacy teacher.

She participated in a GEMS workshop many years ago.

Unit implementation. The unit took the teacher almost four months to complete, from January to April. She used the Terrarium materials three times a week on average, and for a total of 3-4 hours each week. The teacher used everything that was sent to her and did all activities that were included. For the ELL considerations, she only used it once or twice when she had a Russian ELL in her class. The teacher team-taught with another teacher and the other teacher brought in some supplemental books on grasslands and tropical forests during the science lab sessions for the students.

Homework activities were given to students but the general student return rate was about 50% for each activity. She administered the magazine assessments, both pre- and posttest, and conducted critical junctures as well. She used the scoring rubrics provided to score student responses, and she incorporated the scores in assigning students grades for science and to check whether or not the students understood the lessons or not and to see what concepts the students understood.

In the classroom, the teacher typically started with shared reading according to the teacher's guide, and then had students re-read the materials through pair reading (pairing low students with high students). She devoted more *Terrarium Investigations* to science while using Terrarium as she normally would.

Student engagement and motivation. Students enjoyed the Unit and they learned a lot. The teacher noticed an increase in their science vocabulary and knowledge that they didn't have before. Students especially liked the part they learned about earthworms and isopods. They were really engaged in having those live, real-life experiences. The following were some example comments from the students:

> "Ms. ____, I really liked getting the trays ready with the worms, giving them out to the students and giving them the little squirt of water…"
>
> "I really loved seeing all the legs on the isopods."
>
> "I loved holding the worm up to the light to see what was inside of its body."

The teacher found the writing guidelines really good in helping the students to focus. The reading lessons did not help really low kids to improve their skills, they were too hard, while they did help the average to above average students to improve. Students learned to behave like scientists while they were planning, organizing ,and doing their discourse circles. They acted very positively towards the Unit.

Unit quality, usability, and utility. The teacher found the Unit to be of great quality and covered the contents thoroughly. The assessments and the scoring rubrics were time-consuming to use but they provided the teacher with lots of information about the students, especially the pre- and post-test. The teacher was at first overwhelmed by the teacher's guide and then found it useful once she got used to it.

The students did not like the pretest that much because they didn't know most of content, and they found the posttest much easier. They liked the readers and enjoyed reading them to each other and to their kinder-buddies. Readers supported really well what the students were learning.

The teacher thought the unit was pretty good in balancing science and literacy, and that her students were challenged by the contents in both science and literacy.

She used another literacy program in addition to Terrarium, and she would continue to use Terrarium as supplemental for literacy and as a main program for science. She was sure that she would use the materials again next year and would definitely recommend it to other teachers.

Problems in implementation. The teacher was a bit concerned with the timing of lessons and some activities required lots of advanced preparation. She could not follow the suggested timeline, it took much longer to cover the materials. She had some difficulty in getting the different roots in winter time as specified in the Unit. She commented that the materials were not labeled to indicate the uses and purposes. For example, the teacher was not aware that the little apple caps buried in the yellow bin was supposed to be used to cover up the holes to prevent the worms from getting out. There was a report cover and a timer or a stopwatch that was never used. And because of the fire code regulations in Colorado and at the school, it was a challenge to keep track of all the charts on the wall since they would otherwise cover up the whole wall.

Students found it hard to move from the evidence they got from their investigation to generation of an explanation. They needed multiple examples and a lot of scaffolding. Maybe having a little demonstration video of students going through that whole process of creating the investigation would be helpful. The chart worksheet associated with the snail investigations did not have enough space for students to get all the information into it.

Student learning. The teacher was amazed with the huge growth she saw in the students' literacy proficiency, vocabulary knowledge on the science in the literacy test, and their understanding of concepts on the science test and magazine assessments. At the end, student would start to do things without any instruction like labeling and observing.

Most effective for. Students at and above grade level benefited the most from the unit, but overall all students benefited a lot from the Unit.

Factors for successful implementation. The two main factors that contributed to the success were: (1) the teacher received all the materials in the mail, and (2) the Unit was of a very high interest for students.

Recommendations and suggestions.

- Have some examples to show students how to develop a question and make an investigation, etc.
- Have some explanation on how to use the materials.
- Students liked to have materials in the style of Eye-Witness, published by Dorling Kindersley, where there are lots of facts and lots of pictures. Students could either read all the text, or they could skim for titles, it would be up to students what they wanted to read.
- Have readers in a bigger size for shared reading.
- Have readers in different difficulty levels for different students, covering the same contents.
- Have students involved in the preparation stage instead of teachers alone.
- Have the chart/worksheet in a larger size.
- Have some initial training for teachers, something like a video.
- Label the contents in the packaging with information like which lessons the materials went with.

Teacher learning. The teacher learned about earthworms and isopods which she did not know much about before.

Profile for Teacher 5

Classroom background. This teacher used the Terrarium Unit with her 16 3rd graders and 15 4th grade students. Out of her 16 3rd graders, 81% were estimated to be above grade level, 13% at grade level, and 6% below grade level in both science and literacy. The corresponding percentages for the 15 4th graders were 6%, 47%, and 47% respectively in both science and literacy. There were two bilingual students, but they were at grade level. For the six special needs students in the 4th-grade, the teacher gave them extra help in the way of a scaffolding task, or an additional repeat on what's being asked to compensate for their poor processing skills.

Teacher background. This teacher spent six hours in the past 12 months and 18-20 hours in the past three years in science professional development; she attended 4 hours of literary professional development in the past 12 months and 25-30 hours in the past three years. She has a BS degree in biology and has an elementary teaching credential. She ranked herself as "high" in her computer experience/knowledge, in her earth science knowledge, and in her expertise as a literacy teacher, and she ranked herself as an "expert" in her science knowledge.

She participated in at least one GEMS workshop before.

Unit implementation. The unit took the teacher about 10 weeks to complete. She used the Terrarium materials for an hour per day, five days a week. She did all home activities, the magazine assessments, and both the pre- and posttest. This teacher followed everything in the Unit except she didn't use the scoring rubrics. She extended some of the writing tasks and had students hold group discussions so she could evaluate their understanding. She supplemented the Unit with her own readers, some worksheets, graphic organizers, and the software "Inspiration" to build visual concept maps. These extra tasks and supplements were mainly for the 4th graders in her class. The teacher also added in some vocabulary assessments to the Unit.

In the classroom, the teacher typically asked students to do independent reading followed up with students responding to teacher's questions. She used the information from the assessments to know whether the students were getting it or not, whether she needed to go back and repeat or adjust, and whether she should modify her teaching. With Terrarium, the teacher thought that she spent more time on science as she used to.

Student engagement and motivation. Students loved the Unit, they were engaged in learning, and enjoyed learning. They were thrilled with what they were learning. Students really enjoyed working with the organisms. They were very curious, and very much attached to their terrariums. And they were so excited—they would go out to their own garden, "oh my gosh. Can I put this root in the worm bin or in our terrarium?" They incorporated their computer skills in their work to make PowerPoint slides for the curriculum fair at the end of the year.

Unit quality, usability, and utility. The teacher thought the materials were of high quality and easy to use, and the materials prepared the students for the next level well. She had been recommending it to the other teachers. She really appreciated the sequence of the Unit and how everything was built. The background information was very useful to her, though it was a bit wordy. She enjoyed five of the readers and found "Into the Soil" and "My Nature Notebook" marginal. She didn't implement the "My Nature Notebook" in her class because she didn't find it useful, there was a lack of time, and the students had a science notebook already.

The teacher really liked the homework activities because they involved parents in the process so they knew what was going on at school and they could share their experience with their kids. She liked the flip chart very much, it enabled people to go back and refer to it. She liked doing the posttest since she preferred to teach a lot of materials, do the informal assessment, observe, sit down, ask direct questions to the students, and then give the students a test in the end. She found both pre- and posttests helpful to measure progress; and discourse circles insightful to evaluate student understanding and to promote student learning by making new connections and asking new questions.

The teacher thought the unit was pretty good in balancing science and literacy, and that her students were challenged by the contents in both science and literacy. She would definitely use it again as a main if Terrarium aligns with her district curriculum or as a supplement if not.

Teacher problems in implementation. There was no problem.

Student problems in implementation. Students had a difficult time in designing an investigation and doing a science inquiry. They needed a lot more practice and guidance.

Student learning. Students learned how to locate information, ask the right questions, and use information to support their statements. For example, a student would say "wow, how could you say that you don't think there's any kind of decomposition going on in there?" The other would reply "well, because…" and then he or she would have to defend his or her position. This would help students in their persuasive writing.

Most effective for. The unit is effective for all students and students who wanted to learn more learned more.

Factors for successful implementation. Having all the teaching materials all included in the kit contributed to the success greatly.

Recommendations and suggestions.

- Prefer the writing tasks to be more challenging
- Prefer more development on vocabulary
- Like to have more coverage on the application of how to apply the reading skills such as cause and effect, main ideas, fact and opinion, sequencing, etc.

Profile for Teacher 6

Classroom background. This teacher used the Terrarium Unit with her 2nd grade students. Out of the 19 students, 30% of them were estimated to be above grade level, 50% at grade level, and 20% below grade level in science proficiency. For literacy proficiency, the estimates were 10% above, 40% at, and 50% below grade level. There were six English language learners (ELLs), two of them were advanced, three were at the intermediate level, and one was at the beginner level. For the ELLs, the teacher used GLAD and SIDAI to give extra help on vocabulary, to make things more visual, and to have more hands-on activities. She also had two special education students who had severe learning disabilities, but she didn't submit information on them. They were transferred out in the mornings, and a teacher's aid worked with them when they were in the mainstream classroom in the afternoons.

Teacher background. This teacher spent zero hours in the past 12 months and zero hours in the past three years in science professional development; she attended zero hours of literary professional development in the past 12 months and 10 hours in the past three years. She has a BA degree in social studies, a multiple subject and single-subject in social studies teaching credentials besides her C-CLAD. She also took about 75-80 units beyond her BA. The teacher ranked herself as "moderate" in her computer experience/knowledge and in earth science, "high" in expertise as a literacy teacher, and "high" in her science knowledge.

She participated in at least one GEMS workshop before.

Unit implementation. The unit took the teacher four months to complete. She started using the Terrarium materials for an hour per day, five days a week, and after realizing how intensive it was, she increased her time allocation to two hours per day. The teacher used everything that was sent to her and did all activities (except the 1st home activity since they had 10 feet of snow in that period) that were included. For the ELL considerations, she used them whenever she found it fit. She administered the magazine assessments, both pre- and posttest, and conducted critical junctures as well. She used the scoring rubrics to score student responses, but they served as information to her rather than as a basis for grades.

She supplemented the Unit with some of her own assignments—having students write a paragraph for earthworms, for isopods, and for The Reasons for Losing Soil, separately. These assignments were given to make her students talk

with their parents and to see what they had learned in the classroom. The teacher also brought in a lot of songs from the Banana Slug Swing Band for literacy and one video on worms plus a worm book for science.

In the classroom, the teacher used all reading approaches, while most of the time she asked students to do shared reading, and buddy reading for the lower readers. For one student who was struggling with writing, she would type up what he said and had him do the illustration and labeling. For the class, she would use some of the examples from student assessments to demonstrate her expectations.

Student engagement and motivation. The students loved the Unit. They were so excited with the Terrarium Unit, especially with the worms and the isopods. They loved the discourse circles and the presentation activity. Some of the students would even make their parents take them to school when they did not feel well in order not to miss the science lesson.

Unit quality, usability, and utility. The teacher found the Unit to be of great quality, and easy to implement with the step-by-step instruction, e.g. on how to use the readers, how to do the writing, and what techniques to use for literacy, etc. She really liked how the Unit put together reading, writing, and science. The Unit's most useful part was to tie literacy with science by having students read non-fiction and learn to write non-fiction.

She thought the readers were great, especially because they were written right for this unit so they really answered the questions kids had and they were so directed at what they were learning. And the other amazing thing was to have all the materials (worms, plants, etc.) sent over with a phone call.

She found plenty of assessments to use and thought the magazine assessment was kind of unique because it covered more than just vocabulary—students had to read before answering. She found this a really interesting way of assessment. She liked the literacy assessment where students had to read as many words as they could in a minute, and she also liked the vocabulary one because it showed her what kind of vocabulary her students had. She saw a huge jump from the pre- to the posttests on vocabulary. She contributed the growth to the readers and the teaching approaches—teaching the vocabulary and teaching the concepts which were missing in the Shoreline Unit.

The teacher found the assessments useful in measuring student progress and thought that the scoring rubrics were great. The rubrics were very effective in

giving the teachers something concrete to compare the kids' work to. Though it was time-consuming to use and score all assessments, the combination of the assessments let her learn the most about her students, especially from the pre- and posttests.

The teacher thought that the Unit was pretty good in balancing science and literacy, and that her students were challenged by the contents in both science and literacy. The Unit had really high standards and that really pushed the kids. She was amazed how much students learned when they were challenged, it made her think that "oh, I need to raise my level of expectations for them…" She would definitely use it again as a main program if Terrarium aligns with her district curriculum or as a supplement program if not.

Teacher problems in implementation. The most difficult thing for the teacher was to help kids to come up with their investigation questions and summarize what they learned, maybe because the kids were too young.

Students problems in implementation. Students had a difficult time to design an investigation and do a science inquiry. They needed lots more practice and guidance.

Student learning. The teacher was amazed with the huge growth she saw in the students' literacy proficiency, vocabulary knowledge on the science in the literacy test, and their understanding of concepts on the science test and magazine assessments. They also impressed the other students in a higher grade and their teachers with their knowledge and the amount of vocabulary they demonstrated at the presentations. At the end, students would start to do things without any instruction like labeling and observing.

Most effective for. The unit was effective for the basic and advanced students, while the below basic students learned a lot too.

Factors for successful implementation. Having all the teaching materials included in the kit contributed to the success greatly.

Recommendations and suggestions. Prefer to have the guide layout as that of the GEMS guide.

Teacher Learning. The teacher learned to look at reading, writing, and science more as a unit, and the importance of that. She learned a more effective way of teaching the non-fiction writing.

Appendix F:
Terrarium Investigations Coding Summary

Background on Classrooms and Teachers

Three teachers taught 2nd grade, two teachers taught 3rd grade, and one teacher taught a combined 3rd/4th grade class.

Students' level of literacy proficiency.

- Above: One teacher had 10% of their students above grade level, two teachers had 22%-25% of their students above grade level, and two teachers had 40-50% of their students above grade level. For the teacher with a combined 3rd/4th grade class, 81% of the 3rd graders were above grade level and 6% of the 4th grade students were above grade level.

- At: One teacher reported 25% of students at grade level, the other four teachers had between 35%-50% of their students at grade level. The teacher with the grade combined class had 13% of the 3rd grade students at grade level and 47% of the 4th grade students at grade level.

- Below: Four teachers reported 25%-30% of their students below grade level, and one teacher reported 50% of the students below grade level. The teacher with the 3rd/4th grade combined class had 6% of the 3rd grade students below grade level and 47% of the 4th grade students below grade level.

Student's level of science proficiency.

- Above: One teacher reported 9% of their students as above grade level, and four teachers reported between 20%-30% of their students above grade level. The teacher with the combination class of 3rd and 4th graders reported 81 % of the 3rd grade students above grade level and 6% of the 4th grade students above grade level.

- At: One teacher had 22% of their students at grade level. The other four teachers reported between 50%-60% of their students at grade level. The teacher with the grade combined class had 13% of their 3rd grade students at grade level and 47% of 4th grade students at grade level.

- Below: One teacher reported 70% of their students below grade level. Four teachers reported between 10%-25% of their students below grade

level. And the grade combined teacher had 6% of the 3rd grade students below grade level and 47 % of the 4th grade students below grade level.

Amount of hours spent on science professional development in the last year and three years:

- Each teacher spent between 0 and 24 hours in the last 12 months.
- Five teachers spent between 0 and 18 hours in the last three years. One teacher spent 45 hours, and the remaining teacher spent 100 hours in the last three years.

Amount of hours spent on literacy professional development in the last year and three years.

- Each teacher spent between 0 and 20 hours in the last 12 months.
- Two teachers spent only 10 or 12 hours, but the four others spent between 30 and 60 hours.

Highest level of education. Three teachers had a Bachelor's degree, of which one had a certificate and extra units. Two teachers almost had a Master's degree. They only had to finish up 2 more weeks, and one teacher already had a Master's degree.

Teaching credentials.

- Teacher One: K-9 teaching certificate for Illinois
- Teacher Two: National Board Certification, rank one (Kentucky)
- Teacher Three: Early Childhood credential and endorsement for English language acquisition
- Teacher Four: Accredited 2-5 elementary education to teach all subjects. (BA in Elementary Education, Master's Degree in Elementary Reading and Literacy)
- Teacher Five: Multiple subjects, single subject in social studies, Cross-cultural, Language and Academic Development credential
- Teacher Six: Early childhood certificate of completion, standard teaching license for elementary school (Oregon)

Unit Implementation

Days of using the unit in the class.

- Five teachers indicated they used the Terrarium unit every day. One of these teachers had an exception, there was no instruction for four days due to a field trip.

- One teacher mentioned using the unit three days a week on average.

Materials used. In general all teachers used everything.

- Readers: All teachers used the readers.

- Teacher's guide: All teachers used the teacher's guide.

Activities used.

- All six teachers used all types of the assessments: embedded assessments, magazine assessments, pre- and post-assessments, and critical junctures. Some exceptions were: One teacher used only some of the embedded tasks and critical junctures; another teacher did not use every single assessment every time.

- All teachers used the ELL considerations to a certain extent. One would glance over the section and implemented one activity or two, someone else said that she used them sometimes when they were good for everyone. A third teacher used a few considerations for one of her beginning language learners until the student moved out of the class. Teacher Four mentioned using some of the considerations, specifically the ones that had graphic organizers to help her disabled students (speech, emotionally disabled, and learning disabled). The fifth teacher would read over them and use them as needed. The last teacher used them as a reference point to make sure she covered the main ideas or to see if there was a better way to explain the main ideas. But it was more of glancing over than anything else.

- Most teachers used all of the home activities. Two teachers did not use all of the home activities because of the weather.

Reading approach used. Three teachers indicated they used all three reading approaches: independent, paired, and shared. Others mentioned they used a combination of approaches.

- Independent reading: Four teachers mentioned they used some independent reading. One teacher hardly used independent reading, she only used it a couple of times where the kids had to read and share with their parents as their homework. Two other teachers used independent reading because it was instructed in the teacher's guide. The last teacher used independent reading with some students and guided reading with other students at the same time so the books could circulate throughout the classroom. Following is her explanation.

"Because there were only 10 of the readers, that was difficult for me to have everybody doing the readers. I do groups of kids. And so it was difficult for four students to share one book without someone checking out and not really paying attention. So I did some guided reading. But most I did independent so that those books could circulate throughout the room. And then I would go back and say—this is after they had read it once—I would go back and say, "hey wait a minute, let's go back and what was happening here? Who can find it?" And then I would actually make kind of a bit of a game of who can find where it tells me about what adaptation—you know, I would take those questions and I would have them try to dig through and try to find the information. So to me, in a way, that's a little bit of the guided. They did a little bit of buddy. But most of it was independent."

- Paired reading: Five teachers used paired reading. One of the teachers almost used it every time because she did not have enough books for everybody to read on their own. She paired a less able with a more able student and that worked well. One student would be able to help the other student with the vocabulary. Another teacher mentioned using paired reading as it was suggested in the teacher's guide. And a fourth paired her high-reader up with a low-reader. She stated that:

"Oh, also, we have kinder-buddies. And so sometimes they would get an extra reading in then. Reading it to their kinder-buddies. And if it was a low student, and they had trouble with the words, I would tell them to just talk about the pictures or what they remembered about it."

Shared reading. Four teachers used shared reading. Three teachers used it as suggested in the teacher's guide. The fourth teacher used some guided reading.

- Other: One teacher used a combination of the Fluency Oriented Reading Instruction with independent and shared reading. The following described how the reading approaches were used.

"Okay, it depended on the text and what I was hoping to get from the text. But I'm also doing some called Fluency Oriented Reading Instruction and we're doing testing on that.

And that's an established program that when I'm doing Fluency Oriented Reading Instruction, the first time we go through the book I read it out loud. The second time is read along, and the third time is echo reading, and the fourth time is paired reading. And then they're asked to read it at home on occasion. Some of those books I actually used for that. Some of them were easy enough that they did independent or paired reading with me, giving a lot of assistance to the two or three kids that had trouble reading. So I used many different techniques."

Use of assessments.

- One teacher, who said that the assessment was her weakest part of doing the Terrarium Unit, would look at the rubrics, look at the students' work, and then try to make some decisions about whether they understood the key concepts or not through conversation. If there were common errors, a small discussion or mini lesson on those particular errors was given to the whole group, a small group, or an individual. Giving the students a letter grade was done too, since students wanted to have grades and it was required to give grades.

- One teacher used the assessments as instructed in the teacher's guide. She typically walked around the classroom and gave mostly verbal feedback while she walked around the class as the students were doing the work. If she saw that her students had difficulties with a concept, she did another activity to make sure they understood it.

- A third teacher used the basic scoring guide that was given in the teacher's guide. She used the information from the assessments as part of her students' science grade.

- Another teacher stated the following:

"The critical junctions it was more of an observational, and then give extra instruction for who needed it. And the other I used the rubrics to grade, and that was part of my grading of the students. Because unfortunately we do have to collect the grades for the grade book. For the report cards. And then I used the magazine assessments, the first one, I did not take an assessment grade on, but I looked at them to see where my students were. And then the second one, I was looking for growth. So it relied a lot on growth. I also took a literacy grade for the magazine because we worked on restating a question and writing a complete response in other curriculum areas. And I expect the same thing for that, so I used one of my own literacy rubrics to take a literacy grade also for the magazine. And I also graded the pre- and the post-, and that was purely for me to see their grades. That did not go in the grade book. It was just to let me know."

- A different teacher used all the tests as mentioned in the guide, she would score them and compare them. She did not score the embedded assessments.

- The last teacher used the assessment as a way to see if her students were getting it or not. She would go back and repeat the instruction or adjust her teaching as needed.

Student Engagement and Motivation

All teachers said that their students actively took part throughout the whole unit. Following were some examples:

> "Well, the days when we did worm bin observations, we always went way over time, because anything that was hands-on they were just really interested in doing that."

> "The discourse circle thing. That still just freaks me out. It still just surprises me…. They came up with some really good reasons, and almost everybody was able to come to a consensus, even my little stubborn ones gave in eventually…. But because I guess I was expecting so little from that, and I got way more than I expected."

> "And they enjoyed taking care of the organisms. Like squirting the plant terrariums and the worms and the animals, and making sure they had food. The kids brought in food for them."

> "One of them today said, "Ms. ____, I really liked getting the trays ready with the worms, giving them out to the students and giving them the little squirt of water…" So, you know, it was like the stuff with the live critters that really drew them in. "I really loved seeing all the legs on the isopods." And, "I loved holding the worm up to the light to see what was inside of its body."

> "Even though, when we were doing the worm bit, and there was the "ew, oh no, gross!" They were very curious still, and very much attached to their terrariums. And they were so excited-they would go out to their own garden, "oh my gosh. Can I put this root in the worm bin or in our terrarium?"

> "They have such buy-in to these terrariums. I think I told you, we had a scientist visit us—she was doing a presentation on a black widow spider, and she was teaching the kids about habitats, and didn't realize that they were such experts. So we invited her to come by, and she came by and they for half an hour just showed her everything. And used vocabulary. And we're explaining about their adaptations, and explaining their habitats, opened up the worm bin, and then sang her songs about decomposition. And

they were thrilled to have an audience. As they were with their presentations. They were so grown up to be able to go to upper grade kids and do presentations. They still would like to go to more classes. So, we may do that."

Students' reaction. All students loved the Terrarium Unit. Following were some examples:

"Overall, at this level they liked it because there were living things in it, and there were lots of opportunities for interaction between the students, and the materials, and things to do."

"…as a generalization, that's what they like. And because they were doing those kinds of things and the books related to that, that encouraged them to want to understand the books. Because there was a relationship between the books and what they were actually doing, and the books helped them understand what they were doing, as opposed to just reading a book and trying to understand the concepts that you may or may not ever want to know. So I thought they reacted very well and I've gotten good feedback from the parents saying that the kids come home excited about it, and talking about it."

"I didn't have anybody that didn't like it that I know of. They were a little stressed out about doing the presentations, but they did such a good job with them. That was another good thing about it. I've watched fourth and fifth graders give presentations and read it off of a sheet of paper."

"Well, I asked them today, 'would you recommend this to other third graders?' And they were all like, 'Yeah!' One girl said, 'now I know how to design an investigation.' She's one of my big GT kids. 'Now I can design my own investigations at home.' I mean, it was all positive."

"The kids loved them…. And these students, they say, they comment on the fact that my homework is the most fun homework that they've ever received, and they ask for it throughout. 'When are we going to have another homework assignment from you?'...And these are students that do not like to do schoolwork! ...the science block is their favorite time of the day! (laughing) They know that it's going to be a good mixture. They get to communicate with their peers. They know that they really, after the practice of the think/pair shares, and the other discussion-type activities, they know what to expect in those routines, so they automatically do it now. And they know that they're going to get to do hands-on things, and that they're going to get to read about those experiences, also, in real life…the group projects that they had to work on, the students were very excited each day to come in…it was so guided and it's structured where they knew what was expected. I guess that's what, even in those, the

communication activities, the students knew what was expected of them during those activities."

"They loved it. They were so excited when we did the terrariums, when we got the worms, and when we got the isopods. And they loved the discourse circles. They just were so proud of themselves acting like scientists. Their posters, they're so proud of their posters! And when they went to the other classrooms, they just loved it. And I think I told you last time, that I had kids who didn't feel well, and there parents had said well, they didn't want to miss science so I had to bring them today [to school]. That's what you want!"

"They loved it…they were thrilled and wanted to incorporate computer skills, so we had done some PowerPoint slides. So they were engaged and really enjoyed it."

Unit Quality, Usability, and Utility

All six teachers found the curriculum/teacher's guide useful. Some comments teachers gave were:

"I think it would be extremely useful. When it was talking about improving reading skills, again we were doing the Fluency Oriented Reading Instruction. And I would look and blend the two together, but a lot of these things are things that I was already doing as far as reading instruction, but it wasn't linked directly to science…"

"Very useful. I wouldn't have been able to do it without the teacher's guide."

"I thought it was great. At first it was a little difficult for me to figure out how to use because there was so much information in there. And then I found it really valuable because they had the lessons on one side, and it was step by step. And then on the right side, they would have just things to consider for ELL students, or if your kids needed a little bit more, and I thought it was very thorough."

"I really appreciate all the background information. I think that is critical. It's one of the things I like a lot about GEMS."

Here were some concerns:

- One teacher mentioned that the teacher's guide was quite overwhelming.
- Another mentioned that the teacher's guide was a little wordy; she liked more bullet point or key ideas. She mentioned that it would be nice to

have a little bit less text per page. And also to have the vocabulary words with their definitions running down the side.

Readers. In general all teachers found the readers very useful.

- One teacher mentioned that two of the readers ('Into the Soil' and 'My Nature Notebook') were a little marginal, two other readers medium. She really liked five of the readers.

- Another teacher found that the readers supported really well what the students were learning. They liked them and they enjoyed reading them to each other and to their kinder-buddies.

- One teacher liked the shorter readers as opposed to the science book because it was easier for students to go back and re-read them and the readers could be sent home. She would like to purchase more eventually.

Reader reading level. The following were some statements from the teachers about the readers' reading level. The last teacher thought her male students were more motivated by the readers.

> "When we first started the first unit, it's hard for me to tell for sure if it was my perception, or it really was, I thought that they were very difficult for the students….But now I'm seeing about 75 percent or more of the students can read them independently. It's certainly their instructional level. You know, if we go over them once and talk about the concepts and then they can read them pretty well by themselves. So I think it's good."

> "The reading level was good. The content made it where even the high readers had to-that it was still challenging, for the most part, the readers were still challenging to them because it was new vocabulary, a new way of looking at something."

> "The reading level I thought was great for my grade level and my upper level kids. My lower level kids, as long as I had them buddied, they were able to read it. And also, the vocabulary had been introduced so much already that the kids were able to read these long, big words and understand them. I think that's where the unit and the readers went really well together."

> "…and of course, you know, with the ones that are more capable, they had very little difficulty with it. Then on the other hand, the ones who always struggle, struggled. And yet sometimes my little boys who aren't usually the best readers are very familiar with a lot of this because that's where their interest lies. So they were able to probably read better then they would have if they were reading a story."

Reader science level. None of the teachers mentioned anything specific to the science level.

Unit balance. Three teachers thought the unit gave balanced attention to science and literacy. One added some more reading material, and one commented that:

> " I think it did a great job because we had the readers to read, and then we wrote about it, so it was just so well-integrated."

One of the other three teachers mentioned the Unit only met the science goals and a small portion of the literacy goals. This teacher would add a lot of other kinds of reading and used it as a front-loaded for science. Specifically:

> "We would do a guided reading at the level of the student, so we would have smaller groups reading at different levels. And I've always kind of wanted science books at different levels for different topics, so everyone in the class could read and discuss about the same topics and yet everyone in the class wouldn't be reading necessarily the same book."

The other two teachers said it could be heavier on the literacy.

Usefulness of assessments to measure progress. One teacher hadn't scored the posttest yet at the time of the interview and therefore couldn't answer if the assessments measured progress. The other five teachers all stated that the assessments measured growth.

- One teacher specifically said that they would have been extremely useful if they were used the way the teacher's guide said. Unfortunately this teacher didn't have enough time to go back and re-teach. Re-addressing happened, but there was not another assessment to see if the students had mastered it that time.

- Another teacher mentioned progress or mastery even.

- Someone else stated that the student growth was huge.

> "Because there was such a big leap from the beginning where the kids were asked to do things without any instruction like labeling and observing. And then when they were given some instruction and had rubrics, it was amazing to see the growth."

- Another teacher stated:

> "I think that the pre- and post- are helpful. And, you know, in terms of an informal, those discourse circles I thought were insightful because at that point, I could see kids learning how to justify a position, and you could see that they were starting to connect dots. When they had to explain why they thought something was happening, or when, 'what're some new questions that came up for you when we were doing this?' Kind of putting that out there and hearing what was coming from them, you could see new connections and new questions coming, so…."

Usefulness of assessments to measure progress ELL's. None of the teachers mentioned specifically that the assessments and rubrics were useful to measure EL students' progress.

Using the Unit again? All teachers would use the Terrarium Unit again if they had the choice.

One teacher said that it would depend on the district's curriculum and alignment to the standards. It would also depend on the grade level she would be teaching. She would prefer to use the Unit as a main program.

The five other teachers all mentioned wanting to use the science unit as a main unit. Of these five, two mentioned wanting to use the literacy also as a main program. Two others said wanting to use the literacy as a supplement. The third teacher would use the literacy part as a core with branching it out to complete the teaching of the literacy standards (district and state).

Two teachers mentioned the following specifically:

- activity Four, although it is really hard to do it without Activity Three
- all of it

Units liked by students:

- building and planning out the worm bin
- building and planning out the terrariums/isopods
- doing hands-on activities
- presenting posters
- doing investigations

- doing projects at the end of the unit
- reading readers about scientists (The Walk in the Woods)
- doing the discourse circles

Units disliked by students:

- doing the last lesson, Activity 2, the decomposition cards
- taking the pretest
- doing the investigations
- sitting on the floor recording things on charts was too long

Problems During Implementation

- One teacher mentioned that her lower readers had some problems with writing, they needed an extra boost there. One teacher mentioned that the students needed some support to prepare the posters.
- In the first activity (soil) nothing really happened. Not sure if it was the materials or not, but this needed more explanation.

Student Learning

Learning reading concepts and skills well. All teachers said that the students learned the literacy concepts and skills well.

- One teacher said the following:

> "…and I have seen a very big jump in their literacy skills. Every year, our school, one of the benchmark managers they have a STAR—a computerized assessment—and we have to periodically give that assessment throughout the year to track where the students are. And this, and I know you can't compare last year's class to this year's class in a lot of terms, but I have seen a much higher rate of advancement with this class than I have with classes in the past without teaching this unit."

- Another teacher commented that Terrarium had a heavy focus on making predictions, and she would spend less Terrarium time on this the next time. The students were really good at asking questions and note-taking,

and writing informational text. They also had a good understanding now of a table, an index, and a glossary.

- Another stated that the guidelines were really good to help the students focus. The lower students gained vocabulary and the reading really helped the average to above average students.

- Another commented that her students learned the elements of predicting, writing a non-fiction book, using an index, etc. very well.

- Another added that students also learned how to go and get information out of a reader.

Learning reading concepts and skills that did not work well. In general all teachers thought that their students learned the reading concepts and skills well.

- One teacher indicated that the students' writing skills probably needed some improvement.

- Another teacher said that she would spent more time on teaching students to use context clues to figure out words that they were not familiar with.

Learning science concepts and skills well. None of the teachers mentioned that their students did not learn the science concepts and skills well.

- One teacher was very pleased with the level of science concepts the students learned.

- Students were very competent in all elements that were listed on page 16 of the teacher's guide. They could apply their knowledge in a different context.

- It was stated that students' science vocabulary really increased and that they learned the concepts that the students had no knowledge of before.

- Someone else said that the unit was strong. All students had a much higher concept, no matter of what their initial level were.

- One teacher stated the following:

"I thought it was incredible. It was amazing and we did our presentations, that was the only thing we hadn't done when I had my interview with you last time, and we did that like the next day. We went to a lot of different classrooms to present it, even older grade classrooms, and the teachers were just really impressed with the knowledge the kids had,

and the vocabulary. We opened it up for questions, and the kids were able to answer a lot of questions that they had learned in the unit that had nothing to do with their investigation, but they just carried that knowledge over with them. And were able to share it. And I was so proud of them!"

- Some one else thought that her students came away with a great deal. The Unit laid a really nice foundation for the new topics she was teaching, the students were well-prepared.

Students challenged. Most teachers believed their students were challenged by the Unit one way or another.

- One teacher said Terrarium offered a lot of development and opportunities for students at both ends of the spectrum and all of them improved and produced quality work.

Students challenging reading.

- One teacher said that her students learned a lot of critical thinking skills, The vocabulary was challenging to everybody. And almost all writing activities were challenging to everyone. So even though her lowest kids wouldn't have scored as high on the rubrics as the higher kids, they stepped up a little farther than they normally would. They pushed themselves a little harder by the nature of the activities. She didn't have anyone who didn't like the unit.

Students challenging science. One teacher thought her students learned a lot of thinking over and above the science concepts.

Unit Most Effective for

For 2nd graders, one teacher thought that the Unit was most effective for those who were at or above grade level. Another teacher mentioned that is was effective for everyone, especially for her male students.

For 3rd graders, two teachers thought that the Unit was most effective for all students; one teacher stated that the Unit was most effective for all students in science, and most effective for at and above grade level students in reading.

Successful Factors

Reasons why the implementation of the unit was successful:

- A combination of teaching the literacy development skills, and students gaining knowledge that they could apply through the hands-on activities.

- Interesting topics students could relate to. Things they could relate to when they went home, or just went out on the playground. The literacy development provided information that they could use to help them with the activities.

- The fact that it was so hands-on. They had to use their thinking skills all the way through and that built from the beginning to the end. It gradually made them more responsible for the activities towards the end where they had to come up with their own questions and a way to find the answer.

- Having the materials at hand was really nice so you didn't have to go out and get all the things yourself, which is prohibitive for teachers sometimes.

- It was a very high interest topic for the students.

- The excitement in the students. They knew that real life scientists were doing the things they were doing. And then that they were given the opportunities to think on their own, and to make choices within the information. They could choose what went in the terrariums; they could choose what questions they were going to investigate. That choice was really important.

- Scaffolding of concepts.

- By following it step-by-step.

- Having the materials all ready to go.

Recommendations and Suggestions

The following suggestions were made to change the Unit/program to better meet all students' needs:

- Include more support and suggestions for the less able ones in the teacher's guide to support the less able ones a little bit more.

- Have different readers on the same topic for students of different levels. For some students, maybe less text on a page and more picture support with words.

- More support in terms of output in writing and a bit more 'meat' in terms of vocabulary understanding and reading skills.

Suggestions on content changes. The following changes for the unit were suggested by the teachers:

- Writing a fiction story about the critters.

- The adaptation versus evolution section was very unclear.

- One teacher suggested a little demonstration video of students doing an investigation, going through the whole process of creating the investigation, maybe with a separate critter so they did not copy. They could show how to take the evidence and put it into an explanation.

Improvement for materials.

- have a packaging sheet to indicate the contents of usage of the materials

- a reader for each student

- fix the typos

- have a larger worksheet/chart for the snail investigations

- bigger books for shared reading

- copy the style of Eye-Witness where there were lots of facts and lots of pictures

Changes for the teacher's guide. Following were some of the example comments:

- More variation/other strategies for what to do before, during, and after reading.

- Have more options of what to do to support struggling students.

- Include examples of rubrics.

- Teacher's guide should be less wordy. Maybe by putting the most important things in bold, or something else to make it more stick out. Someone else mentioned to use bullet points, key ideas.

- Reminder to use the everyday words and science words

- Training or demonstration video of how to use the unit.

- Teacher's guide: have the extra background information somewhere else bundled together (like the old GEMS guides), so it doesn't interfere with the flow of what you were supposed to do.

Teacher instructional changes as a result of the Terrarium Unit. All teachers changed their regular practices in some way as a result of the Terrarium Unit.

- One teacher changed to teach reading and other literacy skills together. She blended them in more, and that had been very useful.

- One teacher now looked at literacy and science more as a unit.

- One teacher started to use the flipchart more.

- One teacher devoted a lot more time to science that she normally would.

- One teacher used the discourse circles and found it worked well with students with behavior problems:

"The most interesting thing is, this class has been, I don't know how to say this, they have a lot of social issues of getting along with each other and being able to work in a group, even in partners. I have several students who have some real behavior issues and just cannot get along with anybody. But the discourse circle activity just really made a big difference with everybody. And I really thought that that was going to be a bomb. I didn't think that they were going to be able to do that, and now that has come back around in other things that we do. When they have to work in a group, if I say it's just like a discourse circle. You know, you've got to listen to each other, you have to listen to other people's ideas and then you have to come to some kind of an agreement. And that's all I have to say now and they know what they're supposed to be doing."

- She also found that the partner parade had made an impact:

"And another little funny thing happened, I guess. When they do the activity partner parade, I don't know if that was supposed to be in there or if that was something I added in. But I think it was in the beginning. Anyway, I was talking to a little girl about

listening to herself read, and what can you do to help yourself remember what you're reading. And she said it's like partner parade where you have to listen to your partner, only you're listening to yourself the same way, so you have to really focus. And I thought that was interesting that she made that connection. So then after she told that to me, then we all talked about that, so that kind of went on to everybody else."

- The last teacher stated that she had been working on other units of studies and then tried to align other areas, other language arts activities that would support those. She also taught state standards so she would pull state standards that were similar to the things in her other instruction areas.