CRESST REPORT 782

Ellen Osmundson
Yunyun Dai
Joan Herman

YEAR 3 ASK/FOSS EFFICACY STUDY

JANUARY, 2011

The National Center for Research on Evaluation, Standards, and Student Testing

Graduate School of Education & Information Sciences
UCLA | University of California, Los Angeles

Year 3 ASK/FOSS Efficacy Study

CRESST Report 782

Ellen Osmundson, Yunyun Dai, and Joan Herman
CRESST/University of California, Los Angeles

January, 2011

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
300 Charles E. Young Drive North
GSE&IS Building, Box 951522
Los Angeles, CA 90095-1522
(310) 206-1532

The work reported herein was supported by grant number 5387S07-091 from WestEd with funding to the National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

To cite from this report, please use the following as your APA reference: Osmundson, E., Dai, Y., and Herman, J. (2011). *Year 3 ASK/FOSS efficacy study*. (CRESST Report 782). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

TABLE OF CONTENTS

YEAR 3 ASK/FOSS EFFICACY STUDY

Ellen Osmundson, Yunyun Dai, and Joan Herman
CRESST/University of California, Los Angeles

Overview of the Study

This efficacy study was designed to examine the traditional FOSS curriculum (Delta Publishing, Full Option Science System/FOSS, magnetism and electricity, structures of life, and water modules, 2005), and the new ASK/FOSS curriculum (magnetism and electricity, structures of life, and water modules, 2005), a revised version of the original FOSS curriculum, to determine impact on student learning. The enhanced curriculum includes slightly revised instructional materials and a coordinated embedded assessment system, with components developed to support teachers' on-going assessment of student learning and progress towards specified learning goals. The study involves randomly assigned treatment/control groups of 3rd and 4th grade teachers. Year 3 of the project was the study year for ASK/FOSS Cohort 1 (AZ), and the pilot (or practice) year for additional schools and teachers, ASK/FOSS Cohort 2 (AZ, TX and WA). This report presents data from both study cohorts. Table 1 summarizes the data sources, and the type of information provided by each instrument, designed to capture the most critical components of the ASK/FOSS curriculum. In the following sections of the report, analyses and findings are presented based on these data sources and assessment components.

Table 1

ASK/FOSS Study Data Sources and Assessment Components

	I. ASK/FOSS implementation (curriculum and assessments)	II. Assessment strategy use	III. Teacher assessment knowledge	IV. Teacher science content knowledge	V. Teacher analysis of student work
Pre-survey	x	x	x	x	
Content survey				x	x
Observations	x	x	x		
Interviews	x	x	x		x
Teacher logs	x	x			x

Year 3 of CRESST work focused on collecting study year data for Cohort 1 (AZ) including observations, interviews, Teacher Logs and post-content surveys, and the collection of information on Cohort 2 (AZ, TX, and WA) practice year, including demographic information, knowledge, practice, and implementation data using the tools piloted and refined in Year 2 of the project. Student data were collected from both Cohorts, and included pre/post measures for two specific ASK/FOSS modules and state assessment data (language arts, math, and science where available); analyses of student data will be presented in a future report.

In Year 4, the following four research questions will be addressed in the ASK/FOSS project:

1. Are there treatment effects on teachers' knowledge and/or assessment practices?
2. Are there treatment effects on student learning outcomes?
3. To what extent does fidelity of implementation affect student learning outcomes? To what extent do teachers' assessment practices influence student learning outcomes?
4. Under what conditions and for whom does the treatment and/or use of assessment influence student learning outcomes? (e.g., teacher knowledge, student demographics, use of Study Groups)

Teacher Background Information

The pre-survey gathered demographic information on project teachers' education and background, as well as information on entering teachers' instructional practices and approaches to science instruction and assessment. The majority of Cohort 2 teachers completed the survey on-line prior to the Summer Training Institute, while a few teachers completed the survey in hard copy at the Institutes in August 2009. Teacher demographic information from the pre-survey for Cohort 2 (AZ, TX, and WA) and Cohort 1 (AZ) is displayed in Table 2; a summary of all participant demographics is presented in Table 2 as well.

Table 2

Teacher Demographic Information: Cohort 2 (AZ, TX and WA) and Cohort 1 (AZ), Total

	Cohort 2		Cohort 1		Total	
Descriptor	Control *N*=69	Treatment *N*=64	Control *N*=19	Treatment *N*=20	Control *N*=88	Treatment *N*=84
Sex						
Male	3	4	1	0	4	4
Female	66	60	18	20	84	80
Ethnicity						
White	55	43	17	17	72	60
Hispanic/Latino/a	9	13	2	2	11	15
Native American/African American	3	6	0	1	3	7

Descriptor	Cohort 2 Control N=69	Cohort 2 Treatment N=64	Cohort 1 Control N=19	Cohort 1 Treatment N=20	Total Control N=88	Total Treatment N=84
Other	2	2	0	0	2	2
Highest degree received						
Bachelor’s + credential	26	21	5	6	31	27
Bachelor’s + credential + units beyond	12	21	3	4	15	25
Master’s:	17	7	3	5	20	12
Master’s + units beyond	12	13	8	5	20	18
Teaching credential*						
General elementary	63	62	18	17	81	79
General secondary	6	8	1	1	7	8
Special emergency	0	0	2	3	2	3
Multiple subject	2	4	1	1	3	5
Single subject	8	6	2	2	10	8
Bilingual	8	17	4	6	12	23
Administrative	3	2	1	1	4	3
Other: (early childhood, TESOL, guidance, special ed., science endorsement)	12	6	4	5	16	11
Grade level taught						
3rd Grade	49	50	0	0	49	50
4th Grade	17	17	19	20	36	37
Years of experience teaching elementary grades						
Average number of years	10.9	10.7	12.0	8.4	10.9	10.7
Range of years teaching	1-38	2-38	1-32	2-25	1-38	2-38
Years teaching science curriculum unit						
Average number of years	5.0	5.5	3.0	2.6	5	5.53
Range of Years Teaching	1-15	1-18	1-11	2-12	1-15	1-18
Number of science PD hours in the past 2 years						
Average number of hours	12.6	16.5	19.6	21.3	12.6	16.5
Range of hours	0-79	0-100	4-100	2-80	0-79	0-100

*Teachers may hold multiple credentials, so total is greater than number of study teachers. **Note: Cohort 2 pre-survey completion rate: 76% of all teachers.

Demographic information for all teachers in the study is found in the final two columns of Table 2; project participants are primarily white females, with 35% of control teachers holding a bachelor's degree and teaching credential as their highest degree, while 32% of treatment teachers held a bachelor's degree and a credential. Twenty control teachers (or 23%) hold a Master's degree, while twelve treatment teachers (or 14%) have a Master's degree. The majority of teachers in the project possess a general elementary credential. There are slightly more 3rd grade than 4th grade teachers currently involved with the study: 57% of project participants are 3rd grade teachers, with remainder teaching 4th grade. Control and treatment teachers are similar in the average number of years of experience: control teachers have an average of 10.9 years of teaching (range of 1 – 38 years), while treatment teachers average 10.7 years of experience (range of 2 – 38 years). Teachers are similarly experienced with teaching FOSS; control teachers have an average of 5 years of FOSS experience (range of 1 – 15 years), and treatment teachers average 5.53 years of experience teaching FOSS (range 1 – 18 years).

The data in Table 2 similarly suggest few differences between Cohort 1 and Cohort 2 or the treatment and control groups within each, as anticipated due to the study's randomization procedures. The average number of years teaching for Cohort 2 control group was 10.9, with a range of 1–38 years; the Cohort 2 treatment group averaged 10.7 years in the classroom, with a range of 2–25 years of experience. The groups mirrored these trends in their experience teaching FOSS, with an average of 5.0 years for the control group, and 5.5 years for the treatment group.

Cohort 1 and Cohort 2 teachers were similarly comparable in their ethnicity, degrees, and teaching credentials, while there was more variation in years of teaching experience between groups. Cohort 1 treatment teachers reported an average of 12.0 years of experience, while Cohort 2 treatment teachers had 10.9 years of experience. Cohort 1 control teachers reported 8.4 years of experience vs. Cohort 2 control teachers with 10.7 years of teaching experience. Cohort 2 teachers similarly reported more experience teaching FOSS science curriculum than Cohort 1 (average of 5.3 years vs. 2.8 years). Another difference between Cohort 1 and Cohort 2 teachers was the grade level taught: Cohort 1 included all grade 4 teachers, while Cohort 2 involved both 3rd grade teachers (99) and 4th grade (34) teachers (of teachers who completed the pre-survey). Note: The FOSS Modules implemented in this study are designed for use in 3rd and 4th grade classrooms.

The project continuation or completion rate for Cohort 1 was high: of the 39 teachers who began the project in August 2008, 32 teachers (or 82%) remained in the project through its conclusion in June 2010. The project continuation rate for Cohort 2 was similar: 163 teachers completed the practice year, and 134 teachers (or 82%) will continue in the study for the experimental year, 2010-2011. Most teachers who left the project did so because of changes in teaching assignments to different grades or non-project schools (personal communication, M. Tiu, July, 2010).

Pre-Survey ASK/FOSS: Cohort 2 Instructional Practices

Prior to starting the study, Cohort 2 control teachers reported teaching science nearly four times/week (range 1–5 times), for approximately 50 minutes for each lesson, and spent, on average, six weeks to complete each FOSS Module (unit). Treatment teachers similarly reported teaching science four times/week (range 1–5 times), and spent an average of 50 minutes on each lesson. Treatment teachers typically spent weeks six weeks teaching each module.

Teachers were asked to rate their fidelity of implementation for teaching FOSS according to the Teacher Guide: 49% of all respondents reported typically teaching the FOSS curriculum with *moderate changes* to the *scope and sequence* of the modules. Forty eight percent (48%) of both groups of teachers reported typically implementing the FOSS curriculum with *moderate changes* to the *assessments* in the FOSS Modules.

As previously noted in the Year 2 report, teacher self-reports regarding typical implementation of the FOSS Modules highlight the importance of providing teachers with implementation guidelines for the current study, and the rationale for a structured implementation. A number of study teachers reported routinely "supplementing" FOSS science lessons with outside materials and activities based on practical experience working with different curricula. Because fidelity of implementation plays a critical role in establishing the efficacy of the revised FOSS curriculum (ASK) vs. traditional FOSS curriculum, study participants have been requested to implement ASK/FOSS without deviating from the curriculum and Teacher's Guide. To understand fidelity of implementation in this study, data includes both self-report sources (surveys and teacher logs) and third party sources (interviews and observations).

Teacher Assessment Practices

Because the primary change in ASK/FOSS curriculum, relative to the original FOSS curriculum, was the addition of embedded assessments, implementation measures focused on teachers' use of the assessments. To capture teacher assessment practices, four different instruments were employed: (a) self-report use of FOSS assessment activities, (b) self-report of specific FOSS assessment strategies during instruction, (c) self-report of general assessment strategies used during instruction, and (d) Teacher Logs (reported in another section). These instruments were designed to parallel the assessment concepts in the ASK/FOSS curriculum, and to reflect current research and theory on quality classroom assessment. Reliability indices for Cohort 2 teacher self-reported assessment practices scales varied from 0.58 to 0.89 based on Cronbach's alpha (see Table 3).

Table 3

Cohort 2 Reliability Teacher Self-report Assessment Practices

Pre-survey items	Reliability coefficient (Cronbach's alpha)
Establish/communicate learning and assessment goals	0.83
Align assessment with learning goals	0.67
Analyze and interpret student work	0.81
Use assessment to provide information on student learning	0.58
Use of results to plan instruction	0.89

Table 4 displays teachers' typical uses of assessment strategies provided in the ASK/FOSS Modules. Prior to beginning the study, most teachers observed students weekly, made use of student worksheets and response sheets in the FOSS Modules, but made less use of performance assessments. End of module assessments were administered according to the curriculum schedule, that is, approximately once per module at the conclusion of the unit.

Table 4

Pre-Survey: Cohort 2 Self-Report FOSS Assessment Activities

Survey question	Control *N*=69 Mean (*sd*)	Treatment *N*=64 Mean (*sd*)
How often do you use the following FOSS activities as an indication of how well students understand the concepts?		
a. Teacher observations: general	4.5 (0.9)	4.4 (0.9)
b. Teacher observation: specific	4.1 (0.9)	3.9 (0.9)
c. Student sheets (used by students during investigation to organize data)	4.1 (0.7)	4.3 (0.6)
d. Response sheets (used by students to record observations and write explanations)	3.7 (0.9)	3.9 (0.8)
e. Performance assessments	2.6 (1.3)	2.6 (1.1)
f. End-of-module assessments	2.0 (0.7)	2.1 (0.6)

Scale: 1=never or hardly ever, 2= once/Module, 3=every other week, 4=weekly, 5=daily.

Table 5

Pre-Survey: FOSS Assessment Strategies (Cohort 2, Cohort 1, and Total)

Survey question	Cohort 2		Cohort 1		Total	
When teaching a FOSS Module, to what extent do you?	Control *N*=69 Mean (*sd*)	Treatment *N*=64 Mean (*sd*)	Control *N*=19 Mean (*sd*)	Treatment *N*=20 Mean (*sd*)	Control *N*=88 Mean (*sd*)	Treatment *N*=84 Mean (*sd*)
a) Listen/ask questions as students work to gauge their understandings	4.6 (0.5)	4.5 (0.6)	4.7 (0.5)	4.7 (0.5)	4.7 (0.5)	4.6 (0.6)
b) Expect all students to learn the concepts and ideas	4.7 (0.5)	4.6 (0.6)	4.4 (0.6)	4.7 (0.4)	4.6 (0.6)	4.6 (0.5)
c) Analyze and interpret students' ideas based on a developmental framework of how science understandings develop	3.8 (1.2)	3.9 (1.0)	3.9 (0.4)	3.5 (1.2)	3.8 (1.1)	3.8 (1.1)
d) Analyze and interpret whole group discussions for general patterns of understanding	4.2 (0.8)	4.3 (0.7)	4.4 (0.5)	4.3 (0.6)	4.2 (0.8)	4.3 (0.7)
e) Analyze and interpret small group discussions and work for specific student understandings	3.9 (1.0)	4.2 (0.7)	4.3 (0.6)	4.3 (0.7)	4.0 (1.0)	4.2 (0.7)
f) Analyze and interpret individual work and responses for student understandings	4.2 (0.8)	4.4 (0.7)	4.2 (0.5)	4.2 (0.7)	4.2 (0.7)	4.4 (0.7)
g) Use assessments in ways that allow all students to "show what they know"	3.8 (1.2)	4.1 (0.9)	4.1 (0.7)	3.2 (0.8)	3.9 (1.1)	4.1 (0.9)

Scale: 1=hardly ever, 2=occasionally, 3=sometimes, 4=usually, 5=always.

Table 5 displays teacher self-reports of their frequency of FOSS assessment strategies when teaching FOSS. Overall, teachers reported "usually" using general strategies for assessment, such as asking questions to gauge student understanding, analyzing and interpreting whole and small group discussions, and analysis and interpretation of individual work to understand students thinking. Less frequent was the teachers' use of developmental frameworks to understand how students' science understandings were progressing. Teacher ratings for Cohort 2 control and treatment groups varied slightly, but were generally similar. Self-report ratings on general assessment practices were similar for Cohort 1 and Cohort 2, with one question, "use assessments in ways that allow all students to show what they know" showing a statistically significant difference (Cohort 1/treatment group had a lower rating for this item).

Table 6

Teacher Use of General Assessment Strategies (Cohort 2, Cohort 1, and Total)

Survey question	Cohort 2		Cohort 1		Total	
How often do you do the following?	Control *N*=69 Mean (*sd*)	Treatment *N*=64 Mean (*sd*)	Control *N*=19 Mean (*sd*)	Treatment *N*=20 Mean (*sd*)	Control *N*=88 Mean (*sd*)	Treatment *N*=84 Mean (*sd*)
a) Set **daily learning** goals for student learning	4.4 0.9	4.2 0.9	4.5 0.7	4.6 0.7	4.4 0.8	4.3 0.9
b) Set **unit** goals for student learning	4.4 0.8	4.3 1.0	4.4 1.1	4.6 0.8	4.4 0.8	4.3 1.0
c) Communicate learning goals to students	4.5 0.7	4.4 0.7	4.5 0.6	4.6 0.6	4.5 0.7	4.5 0.7
d) Make sure your assessments are aligned with your learning goals	4.6 0.7	4.4 0.8	4.6 0.6	4.6 0.6	4.6 0.7	4.4 0.8
e) Assess students' knowledge prior to introducing a new module	3.8 1.2	3.8 1.1	3.3* 1.1	3.6 1.1	3.7 1.2	3.7 1.1
f) Use multiple assessment methods to gauge learning	4 1.0	3.9 1.1	4.3 0.9	4.1 0.8	4.0 1.0	3.9 1.0
g) Coordinate items on daily, weekly and unit assessments to gauge how student understandings are developing	4.1 1.0	4.1 1.0	3.7 1.1	4.1 0.8	4.0 1.0	4.0 .9
h) Use assessment results as a basis for evaluating student progress toward learning goals	4.2 0.9	4.2 0.9	4.2 0.5	4.6 0.6	4.2 0.8	4.2 0.8

Scale: 1=hardly ever, 2=occasionally, 3=sometimes, 4=usually, 5=always.

In Table 6, teachers responded to questions about their use of *general* assessment strategies (not specific to FOSS). Again, most teachers in both groups reported "usually" engaging in a variety of assessment tasks, including goal-setting, communication of learning goals to students, aligning assessments with learning goals, coordinating items from different time points to gauge student learning, and using assessment results as a way to evaluate student progress towards learning goals. Control and treatment teachers responded in approximately the same manner. Assessing students' prior knowledge was reported as occurring only "occasionally" by most teachers. Again, teachers in Cohorts 1 and 2 reported similar use of general assessment strategies. The only statistically significant item was assessing students' knowledge prior to introducing a new module; Cohort 1 control teachers self-report was significantly lower than for Cohort 2 control and treatment teachers.

The self-report measures indicate that teachers, prior to the ASK/FOSS study, used both general and FOSS specific assessment strategies on a regular basis.

Teacher Content Knowledge

To capture teacher science content knowledge, we used two measures; self-report questions on the pre-survey (see Table 7), and a content survey that included concepts and tasks from the ASK/FOSS Magnetism and Electricity Module (see Table 8, and Appendix A). All teachers in the study teach the Magnetism and Electricity Module, but the second ASK/FOSS unit taught varies according to district and state requirements (the Water Module or Structures of Life Module). Based on available resources, a teacher content survey was developed only for the Magnetism and Electricity Module.

Teacher Content Knowledge: Self-report Survey Results

Teachers were asked to assess their qualifications to teach the targeted concepts in each of the ASK/FOSS modules in the study. The scores are an aggregate of teacher ratings on the primary concepts comprising each module. For example, the Magnetism and Electricity score is the aggregated average of the concepts included in the Magnetism and Electricity Module – i.e., electricity, magnetism, magnetic forces, electrical circuits, and how electricity in magnets produces magnetic effects. Table 7 displays the scale reliabilities for teacher self-reports in each module area, which reveal a high degree of consistency between concepts.

Table 7

Scale Reliability: Teacher Self-Report of Qualifications to Teach Content

Content area	Scale reliability
Magnetism and Electricity	0.95
Water	0.93
Structure of Life	0.94

Teachers' self-reports of their perceived qualifications to teach the concepts in the ASK/FOSS Modules are reported in Table 8. In general, teachers rated themselves slightly higher than "somewhat qualified" to teach the concepts in the ASK/FOSS curriculum.

Table 8

Self-Reported Teacher Science Content Knowledge (Cohort 2, Cohort 1, and Total)

Survey question	Cohort 2		Cohort 1		Total	
How well qualified do you feel to teach 4th grade students about the following topics?	Control *N*=69 Mean (*sd*)	Treatment *N*=64 Mean (*sd*)	Control *N*=19 Mean (*sd*)	Treatment *N*=20 Mean (*sd*)	Control *N*=88 Mean (*sd*)	Treatment *N*=84 Mean (*sd*)
a) Magnetism and Electricity (magnetic forces, electrical circuits, how electricity in circuits can produce magnetic effects)	3.8 0.8	3.7 1.1	3.6 0.9	3.3 0.9	3.8 0.8	3.6 1.1
b) Water (water cycle, properties of water, nature of Earth materials, how water interacts with Earth's crust, oceans and atmosphere)	4.1 0.6	4.2 0.9	4.1 1.1	3.5 0.9	4.3 0.7	4.0 0.9
c) Structures of Life (seeds, life cycle of a plant, animal habitats, how an organism's structures help it survive in its habitat)	4.4 0.6	3.9 0.9	3.9 1.0	3.4 1.0	4.3 0.7	3.9 0.9

Scale: 1=not at all qualified, 3=somewhat qualified, 5=very qualified.

Overall, project teachers reported less confidence in their capacity to teach the concepts in the Magnetism and Electricity Module, and reported greater confidence in their qualifications to teach concepts associated with the Water Module and the Structures of Life Module. For both Cohort 2 and Cohort 1 and control and treatment groups, teachers reported themselves least confident in the topics associated with the Magnetism and Electricity Module. These patterns are consistent with previous research on elementary science teacher knowledge – teachers report more knowledge and confidence in teaching life science topics than physical science (e.g., magnetism and electricity; White & Tisher, 1986). Cohort 1 treatment teachers reported themselves significantly less confident in their qualifications to teach all three of the topic areas queried.

Table 9 presents correlations between teachers' confidence in each of the three module topic areas (Magnetism and Electricity, Water and Structures of Life). Correlations show positive relationships among the three content areas – that is, teachers who feel confident in one content area are likely to feel confident in another, with relationships between the concepts in the Water and Structures of Life Modules stronger than the relationship with the Magnetism and Electricity Module concepts.

Table 9

Correlation Matrix: Cohort 2 Teacher Self-Report of Qualifications to Teach Content

Module	Magnetism and Electricity	Water	Structures of Life
Magnetism and Electricity	1.000 -- 127		
Water	0.542 <.0001 123	1.000 -- 124	
Structures of Life	0.551 <.0001 100	0.768 <.0001 100	1.000 -- 101

Note: Pearson Correlation Coefficients Prob > |r| under H0.
Rho=0 Number of Observations.

Teacher Content Knowledge Survey

In addition to teachers' self-reports on how well qualified they felt to teach specific science concepts, teachers completed a content survey to capture their conceptual knowledge of magnetism and electricity. Teachers' knowledge and expertise in recognizing students' conceptual understandings and conceptual challenges (alternative conceptions) was also assessed. Additionally, we captured teachers' capacity to accurately analyze student responses, and provide recommendations on "instructional next steps" to support student learning, a critical component in effective classroom assessment and instruction. Content surveys were administered in hard copy at the start of the project as pre-test measures of teacher knowledge (content, pedagogical and pedagogical content). The same content survey is administered at the end of the study year, as a post measure for examining study effects on teacher knowledge. Cohort 1's pre-test occurred in August 2008 (see prior annual report), and for Cohort 2, the pre-test was administered during Summer Professional Development session in August 2009. Cohort 1 teachers' post-test was administered during a special convening of the group in May 2010. For both cohorts, participants were allotted one hour to complete the survey, but could request additional time if needed.

Teacher Content Knowledge: Multiple Choice Items

Section 1 of the Teacher Content Survey contained 29 items culled from the FOSS Magnetism and Electricity Module, as well as additional questions from National Assessment of Educational Progress (NAEP) and Trends in International Mathematics and Science Study (TIMSS) 4^{th} grade assessments on magnetism and electricity (Olson, Martin, & Mullis, 2008), and other studies on students' understandings of electricity and magnetism (Heller et. al, 2004).

Three primary conceptual areas, aligned with the ASK/FOSS curriculum, were assessed in the content survey—magnetism, electricity and electromagnetism. These items were reviewed and revised by two content experts from the Lawrence Hall of Science for scientific accuracy and alignment with the ASK/FOSS curriculum; simultaneously a scoring/coding guide was developed and reviewed by content experts to guide the analysis and interpretation of the teacher responses. The overall reliability for the Cohort 2 pre-study content survey scores is 0.66; the overall reliability increased to 0.73 with the removal of four items (2.14, 2.23b, 2.31, and 2.46), identified by their low point-biserial correlations with other items (see Appendix B for a copy of the instrument and specific items). Removal of the items increased the overall reliability of the measure.

As Table 10 shows, Cohort 2 teacher pre-test scores on the multiple choice questions ranged from a low of 24% correct (7/29) to 97% correct (28/29), with a median of 76% correct (22.3 of 29 items). There were no significant differences between treatment and control teachers on the pre-test; scores were similar for Cohort 2 teachers across the three states in the study.

Table 10

Cohort 2 Teacher Content Survey Scores: Pre-test M & E Multiple Choice and Completion Items

	Analysis Variable: Total MC Content Survey Score					
Cohort 2 by State	Mean	Std Dev	Median	Mode	Range	*N*
AZ						
Control	21.9	2.9	22	19	19-26	9
Treatment	22.4	2.0	22	21	19-25	9
TX						
Control	22.3	2.9	22.5	20	13-27	58
Treatment	21.7	3.2	22	20	15-28	53
WA						
Control	20.7	6.3	22	26	7-28	13
Treatment	21.6	2.9	21	21	16-27	14

Cohort 1: Multiple-Choice Post-Content Survey Results

The results presented in Tables 11 and 12 display Cohort 1 teachers' performance on the multiple-choice post test and compares scores before and after participating in the study (i.e., after teaching the ASK/FOSS Magnetism and Electricity unit for two consecutive years). Results show that control and treatment teachers made gains in all areas of the content survey after two years of study participation. For the control group, the difference in pre/post scores is statistically significant at the .05 level for three of the four scales (magnetism and two electricity concepts). For the treatment group, the difference in the pre/post scores is statistically significant for *all* four scales, at the .05 level. For both groups, scores increased most pre/post for items relating to electromagnetism.

Table 11

Cohort 1 Pre/Post Magnetism and Electricity Content Survey Scores: Control Teachers

	Pre/Post multiple choice scores control teachers						
Investigation	*N*	Pre	Post	Pre/post	DF	t value	Pr >|t|
Magnetism	11	0.69	0.83	0.14	10	1.85	0.093
Electricity 1	11	0.68	0.88	0.20	10	4.9	0.001
Electricity 2	11	0.65	0.87	0.22	10	3.36	0.007
Electromagnetism	11	0.50	0.82	0.32	10	1.64	0.1319

Table 12

Cohort 1 Pre/Post Magnetism and Electricity Content Survey Scores: Treatment Teachers

	Pre/Post multiple choice scores treatment teachers						
Investigation	*N*	Pre	Post	Pre/post	DF	t value	Pr >|t|
Magnetism	13	0.75	0.96	0.21	12	5.45	0.000
Electricity 1	13	0.68	0.90	0.22	12	5.42	0.000
Electricity 2	13	0.64	0.90	0.26	12	4.98	0.000
Electromagnetism	13	0.35	0.92	0.58	12	6.04	<.0001

Note: unique teacher IDs were used to match teachers' pre/post scores. Not all teachers completed pre/post content surveys, due to scheduling conflicts, hence the lower number of teacher scores reported than study participants.

In general, we did not find differences in teachers' post-test knowledge as a function of treatment condition (control vs. treatment). One exception was the results for the first subscale (magnetism): when controlling for pre-test performance, treatment teachers outperformed control teachers on the multiple choice post-test items on magnetism.

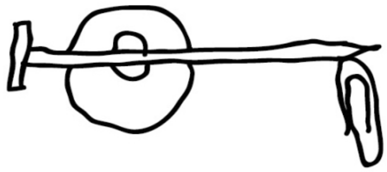
Teacher Content Knowledge: Analysis and Interpretation of Student Responses

Section 2 of the Teacher Content Survey focused on eliciting teachers' capacity to analyze and interpret student responses, a proxy for pedagogical content knowledge. This knowledge helps teachers understand student thinking, and to recognize how student understandings typically progress, and what alternative conceptions students may hold. The structure of these questions was as follows:

5. Teachers answered an open-ended content question;
6. Student responses to the same question were provided, and teachers were asked to analyze and interpret the student responses;
7. Teachers were asked to indicate what the student knows, and what the student needs to learn to progress.

Figure 1 shows a sample item from the Magnetism and Electricity Module that follows the teacher content survey sequence described above.

1.22 Anne is investigating objects and magnets. She made this observation in her science journal.



"I was surprised! A nail was stuck to the magnet. When I accidentally touched the nail to a paper clip, the paper clip stuck to the nail. I wonder why that happened?"

a. Explain to Anne why the paper clip stuck to the nail. Use diagrams or pictures if necessary.

Anne and her friend were asked by her teacher why they thought the paper clip stuck to the nail. Here are their responses to the question:

Anne's response: The paper clip turned into a magnet too.

Anne's friend's response: The nail gets stuck on the magnet, and the nail turns into a magnet, so the paper clip can stick on the nail.

b. What inferences can you draw about the students' understanding of magnetism and electricity? What do these students know? What do these students not know/need to learn?

c. If these students were in your class, what would you do next in your instruction to help the students learning progress?

Figure 1: Teacher Content Survey: Magnetism and Electricity Module.

Scores for this portion of the content survey (Figure 1, part b) were based on a 3-point scale, derived from expert ratings of teacher responses. This coding guide is a revised version from the one presented in the Year 2 report; the revised guide was better able to capture different levels of teacher understanding.

Table 13

ASK/FOSS Teacher Analysis and Interpretation of Student Work Coding Guide

Score	Description
3	Complete response, scientifically accurate identification of student level/s of understanding. Specifically, the teacher response: • Identifies at least one concept of student(s) understanding AND at least one concept that student(s) do not understand. • Differentiates between Student 1 and Student 2 (where applicable). • Does not contain any scientific or inferential inaccuracies. • Correctly identifies scientific concepts in student response, using scientific language. ***Example:*** "Both of these students understand that the force field created by the magnet can extend beyond the actual magnet. Anne's conclusion is correct, and she can test that idea by seeing if the paperclip can attract something else (like another paperclip). The nail is acting as a temporary magnet, and so is the paper clip. Anne's friend appears to understand induced magnetism, and temporary magnets.
2	Partial response, mostly scientifically accurate identification of student level/s of understanding. Specifically, the teacher response: • Identifies at least one area of student(s) understanding AND at least one area that student(s) do/es not understand • Does not necessarily differentiate between Student 1 and Student 2. • Does not contain any inferential or scientific inaccuracies. • Recognizes general scientific concepts in student responses, using general language. ***Example:*** "Anne is not understanding what happens to create a temporary magnet since she says the paper clip turned into a magnet. Anne's friend is closer to what's happening, but doesn't use correct language.
1	Minimal response, minimal level of accuracy identification of student level of understanding. Specifically, the teacher response: • Identifies student understanding OR area that student does not understand in broad, vague or general terms. • Response does not differentiate between students. • Contains inferential or scientific inaccuracy. ***Example:*** "They are both correct. The friend's response is what happens, and Anne's carries it a step further."
0	No response or response that indicates teacher doesn't understand the student response. ***Example:*** "I'm not sure what the student is thinking." OR "Look in the Teacher's Manual for that information."

Due to the large number of teachers in Cohort 2, scoring of Cohort 2 pre-content survey open-ended items, is currently in progress, and will be presented in a future report.

Scores for the Cohort 1 post-teacher content survey for analysis and interpretation are presented in another section below, along with open-ended content and next instructional steps scores.

Teacher Content Survey: Next-Steps for Instruction

In the final section of the teacher content survey, based on student responses and teacher analysis, and interpretation of those responses, teachers were asked to provide information on "next steps" in instruction to help students increase their understandings of magnetism and electricity and to address students' alternative conceptions. A revised 3-point scale, devised by content experts based on teacher responses and then reviewed by content experts (see Table 14), was used to score teacher responses. This coding guide was similarly revised from the original guide presented in the Year 2 report.

Table 14

ASK/FOSS Teacher Instructional Next Steps

Score	Description
3	Detailed, content-specific next instructional steps indicated. Response takes into consideration students' current level of understanding. Response cites a specific strategy that teaches key content AND accommodates both/each students' needs. ***Example*** "Next I would have Anne and her group test whether another object, like a paper clip, is attracted to the paperclip. This would help Anne understand how temporary magnets work. I would also have them remove the nail and see how long the induced magnetism remains. Anne's friend can try the same idea, and maybe use iron filings to see the magnetic force field.
2	General, content-general instructional next steps indicated. Response alludes to "general" level of student understanding. Cites a general strategy that teaches content relevant to students' needs. ***Example*** "I would ask them to explain what a magnetic field is, and introduce them to an electromagnet.
1	Broad, vague instructional next steps indicated. Response sites area/s of focus, but does not identify specific strategy OR cites general strategy with no detail. Suggests general review, redo the investigation, etc., with no detail. Response does not take into consideration students' level of understanding. ***Example*** "I would have them work together to see what they can discover."
0	No response or response that indicates teacher doesn't understand the student response. ***Example*** "I'm not sure what I would do next in instruction." OR "Teach myself."

Teacher Performance on Open-Ended Content Survey Scales

Open-ended items on the teacher content survey were scored by three science educators (including one researcher), specifically trained on the scoring rubric. Educators were familiar with the FOSS Modules, and had an average of eight years of teaching experience. From the total sample of fifty-six pre/post teacher content surveys, twenty-seven percent (15 surveys, 21 questions/survey) were double coded to establish the reliability of scoring. Pre and post surveys were mixed to avoid scoring bias, based on administration time. Coding differences were discussed and resolved. Responses to each of the three item types were aggregated across the specific topic areas of magnetism, electricity, and electro-magnetism to create separate subscales for content, analysis and interpretation, and next instructional steps. Reliability of scoring, based on double scored responses ranged between 76% agreement to 96% agreement and is shown in Tables 15 and 16.

Table 15

Cohort 1: Pre-survey Inter-rater Reliability, Open-ended Responses

Comparison rater	Rater 1	Rater 2	Rater 3
1	1.00	0.96	0.90
		<.0001	<.0001
2	0.96	1.00	0.86
	<.0001		<.0001
3	0.90	0.86	1.00
	<.0001	<.0001	

Note: Pearson Correlation Coefficients, *N* = 63.
Prob > |r| under H0: Rho=0.

Table 16

Cohort 1: Post-survey, Inter-rater Reliability, Open-ended Responses

Comparison rater	Rater 1	Rater 2	Rater 3
1	1.00	0.86	0.91
		<.0001	<.0001
2	0.86	1.00	0.76
	<.0001		<.0001
3	0.91	0.76	1.00
	<.0001	<.0001	

Note: Pearson Correlation Coefficients, *N* = 126.
Prob > |r| under H0: Rho=0.

Table 17 displays score reliabilities for the pre and post content surveys. Results show reasonable reliability for the analysis and interpretation and next step subscales, particularly given the small number of items constituting each. Scores for the content knowledge questions were less reliable than the other two areas, which may be in part due to the small number of content items (a total of seven items).

Table 17

Cohort 1 Score Reliabilities for Open Ended Items on Content Survey

Items	Pre	Post
Content Knowledge	0.51	0.48
Analysis and Interpretation	0.73	0.81
Next Instructional Steps	0.79	0.84

Differences in pre/post content survey scores for Cohort 1, shown in Table 18, reveals that scores for both groups improved after teaching the Magnetism and Electricity Module for two consecutive years. While pre-post differences consistently favor the treatment group, there were no statistically significant differences between control and treatment teacher scores, without controlling for the pre-survey score. For both groups, the largest gain was in the area of instructional next steps, suggesting that project teachers had more knowledge of curriculum and were thus able to identify instructional next steps based on curriculum familiarity, as well as an increased capacity to figure out what to "do next" in instruction to help students' learning progress.

Table 18

Cohort 1: Pre/Post Content Survey Scores, Open-Ended Items

Items	Control *N*=13 Pre/post	Control change	Treatment *N*=16 Pre/post	Treatment change
Content*	4.1/5.4	+1.3	4.4/6.6	+2.0
Analysis and Interpretation**	6.5/10.8	+4.3	7.8/14.7	+6.9
Instructional Next Steps***	6.0/11.4	+5.4	7.6/14.7	+7.1

*Note: scale=1 (correct), 0 (incorrect), 7 possible points. **Note: scale range 0 – 3 (see scales above), 21 possible points. ***Note: scale range 0 – 3 (see descriptions above), 21 possible points.

To further investigate the relationship between pre- and post- survey scores for treatment and control teachers, we utilized a regression analysis, using the pre-survey score for each question type as a covariate. For all three areas – content, analysis and interpretation, and instructional next steps, treatment teachers outperformed control teachers. These data indicate a positive, statistically significant treatment effect on treatment teachers' post-survey scores. Table 19 displays results of the regression analyses.

Table 19

Cohort 1 Regression Analysis for Pre/Post Content Survey

Variable	DF	Parameter estimate	Standard error	*t* Value	Pr > |t|
Post content					
Intercept	1	5.15	0.58	8.96	<.0001
Pre content	1	0.06	0.12	0.46	0.65
Treatment	1	**1.20**	0.39	3.06	0.01
Post analysis & interpretation					
Intercept	1	8.77	1.44	6.11	<.0001
Pre analysis & interpretation	1	0.31	0.18	1.73	0.10
Treatment	1	**3.50**	1.20	2.92	0.01
Post next step					
Intercept	1	9.16	1.46	6.27	<.0001
Pre next step	1	0.37	0.19	1.98	0.06
Treatment	1	**2.73**	1.31	2.08	0.05

Fidelity of ASK/FOSS Implementation

To gauge implementation fidelity of the ASK/FOSS Modules, we used three conceptually aligned measures: (a) classroom observations tied to the ASK/FOSS curriculum, (b) interview questions aligned with the ASK/FOSS curriculum, and (c) teacher logs designed to reflect ASK/FOSS concepts and strategies. Each of the measures contained concepts that were aligned with the curriculum and designed to provide fidelity of implementation of the ASK/FOSS curriculum. We were interested in capturing fidelity of implementation of FOSS to understand teachers' use of assessments in the FOSS curriculum as well as the quality of teachers' tool use. Because each tool was designed to capture the same components in the FOSS assessment system, we had a way to validate, in a sense, teachers' implementation of FOSS and accompanying assessments from different perspectives. Teacher logs were self-reports of assessment practices and assessment use on a weekly basis, while observations captured teachers' assessment use "in action" on a single day. Finally, interviews provided teachers with an opportunity to reflect on and describe their thinking, understanding, and rationale behind specific decisions regarding their instructional and assessment practices during implementation of their second FOSS module. Treatment teachers were tasked with implementing an additional six assessment components (also included in Table 20), as part of the ASK embedded assessment system.

Table 20

FOSS Assessment Components

Data Collection Tool	Interviews	Observations	Logs
Assessment components			
1. Analyzed work in student notebook	x	x	x
2. Analyzed work on student responses sheets	x	x	x
3. Recorded observations of students during class	x	x	x
4. Analyzed student work for patterns and trends	x	x	x
5. Analyzed observations for patterns and trends	x	x	x
6. Planned further instruction based on analysis of student work	x	x	x
7. Checked on student understandings at the end of a lesson on investigation	x	x	x
8. Engaged students in self-assessment of science learning	x	x	x
Treatment specific assessment components			
9. Administered an I-Check Assessment	x	x	x
10. Used coding guides in I-Check Folio to code I-Check items	x	x	x
11. Recorded I-Check codes on "summary coding sheets"	x	x	x
12. Conducted student self-assessment sessions based on I-Check analysis	x	x	x
13. Checked students' reflections after self-assessment	x	x	x
14. Used a "Next-Step Strategy" based on self-assessment sessions	x	x	x

In the sections that follow we summarize data from the observations, interviews, and log data. We conclude with an examination of the inter-correlations among the measures, both as a validity check on each tool, and to inform decisions on the composition of implementation measures for next year's analysis of the effect of treatment on implementation. Further, these analyses also provide information on the relationships between fidelity of implementation on student learning and factors that may influence implementation and effects.

Classroom Observations

Classroom observations were conducted in selected Cohort 1 classrooms in November and December, 2009. The four treatment schools (nine teachers) and two control schools (four teachers) represented a range of school and student demographics, as well as range of teacher knowledge and experience with FOSS and teaching. The classroom observation protocol, as noted earlier, paralleled the core aspects of FOSS and ASK implementation, focusing particularly on fidelity of FOSS implementation and teachers' use of assessment in both treatment and control classrooms.

Data from classroom observations provided background and context for the study, as well as information on fidelity of ASK/FOSS implementation. Additionally, the observations served as a reference point for follow-up interviews regarding teachers' FOSS implementation, including instruction and assessment practices.

Observations occurred during Investigation 4 of the Magnetism and Electricity Module, based on recommendations from Lawrence Hall of Science curriculum developers and professional development experts. Module 4 focuses on connecting the concepts of magnetism and electricity by introducing students to electromagnets. Investigation 4.1 begins by asking students if they have ever seen a "junkyard magnet" pick up a car and move it from one place to another; this question served to activate prior knowledge, and remind students of what they'd recently studied (magnets and electrical circuits). Students were provided with materials – a D-cell, wires, circuit switch, and a metal rivet – to create their own junkyard magnet, without specific instructions on how to do so. In previous investigations, students experimented with magnets and constructed circuits. Investigation 4.1 brings together the concepts of magnetism and electricity together by asking students to apply their knowledge of magnetism and electricity to a new situation.

Note: for instances where specific ASK/FOSS assessments were not in use during classroom instruction, follow-up questions subsequent to the observation were asked to further probe teachers' use of specific assessments and strategies. For example, in treatment classes, end-of-investigation assessments (I-Checks) were not administered during the observation, but teachers showed student I-Checks and the accompanying coding sheets to researchers at the conclusion of the observation.

Qualitative findings in treatment and comparison observations. The qualitative analyses provide a contextual background of the ASK/FOSS implementation in Cohort 1 classrooms. In all of the observed classrooms (12 total), students worked quickly and with enthusiasm to try out various approaches to building or creating "a magnet that can be turned on and off" (Investigation 4.1 title). In the observed classrooms, after approximately 10 minutes of trial and error, one or more groups of students successfully created an electromagnet, and were able to use the "junkyard magnet" to transfer metal washers from one cup to another. In swift succession, all other groups were able to successfully create an electromagnet. Lessons typically concluded with teachers introducing the term "electromagnet – a magnet that can be turned on and off."

In general, classroom observations revealed that Cohort 1 teachers were implementing the ASK/FOSS curriculum with a seemingly high degree of fidelity, according to the “full implementation guidelines” established by the developers and the study guidelines (see Appendix B for a copy of the guidelines). Classrooms, both treatment and control, were rich with science curriculum materials and evidence of work/learning in progress. Posters, word and concept charts, and student work were clearly displayed and referenced during instruction. At the beginning of science instruction, most teachers began their lessons by presenting an overview of the learning goals for the day, and asked students to review concepts and terms from previous or on-going investigations. During science instruction, study teachers actively interacted with students as students carried out the investigations. Teachers asked questions for clarification of student thinking and understanding, added concepts and words to appropriate charts, and observed students in small groups and whole class discussions as they worked to complete the tasks in the investigations. One observed teacher used a formal observation sheet to record information during the course of instruction, while other teachers relied on more informal record-keeping to make notes on student work and progress (e.g., post-it notes in the Teacher’s Manual). Students and teachers appeared to be familiar with the roles and approaches to hands-on science learning, as evidenced by their familiarity with the materials, and strategies for working in small groups, recording information, and sharing findings and observations at the conclusion of lessons.

In all study classrooms (control and treatment), teachers reported that students were enthusiastic and motivated by science instruction, often citing it as “the highlight of the day” for students and teachers alike. One teacher commented that she had initially been very hesitant to allow her students to do science and “investigate and try things on their own.” but saw how interested and motivated students were by the hands-on investigations, and now uses the strategies and assessment techniques she learned from participating in the ASK/FOSS study in other content areas.

Qualitative findings on Cohort 1 treatment teachers’ use of assessments. All observed treatment teachers in Cohort 1 (ASK/FOSS) collected student notebooks at the conclusion of the lessons; most reviewed these notebooks informally during the course of instruction. This year, the second year of implementation for Cohort 1, all observed teachers reported analyzing and interpreting student notebooks at the conclusion of the day as directed by the ASK/FOSS guidelines. Observed teachers reported finding value and important information in the notebooks, and most (8 of the 9 observed) reported using the “what to look for” protocol, a coding guide of sorts, for examining student work at the conclusion of each lesson. Teachers reported using a variety of methods to provide written feedback to students, with most relying on post-it notes and/or other informal methods of relaying written information to students. All observed treatment teachers reported *avoiding marking directly* on students work, and instead used post-it notes or other feedback formats, citing the importance of students seeing their notebooks as their own work and thinking in science, without teacher comments.

In treatment classrooms, the end-of-investigation assessments (I-Checks) were in evidence. All observed treatment teachers reported using the I-Checks as a way to gauge student learning and progress, and found value in the information provided by these assessments. A number of treatment teachers commented on the alignment between the content of the investigations and the I-Checks, noting that the I-Checks provided clear evidence of how well they as teachers had taught the concepts, and how well students had learned the concepts and appropriated the language of instruction. Treatment teachers also reported more skill and ease in using the coding guides to code (score) student work on the I-Checks than during the previous (practice) year of the study.

This year, in contrast to the practice year, most treatment teachers reported engaging students in self-reflection sessions after the I-Checks had been coded. Rather than simply handing back the I-Checks to students and having them correct their errors, teachers described scaffolding subsequent instruction, and making use of the "next instructional steps" to support student learning. For example, in one observed classroom, the teacher selected three questions that were most problematic for students from the I-Check. She used the "critical competitor" strategy to present two different responses to these three challenging questions, and asked students to work with a partner to select the most "complete and scientifically correct" response. Students then shared their thinking, resulting in a rich and conceptually driven discussion about the challenging items. Teachers commented that they tended to use two or three of the strategies for the self-reflection sessions, rather than try new or different approaches from the list of ten possible strategies, because they found that students responded well to certain strategies, and they were sometimes hesitant to "change what was working well."

Interviews that immediately followed classroom observations revealed that, for both treatment and control teachers, the second year of implementation was smoother, more seamless, and more manageable than the first year (practice year), even in instances where teachers reported substantial increases in the number of students in their classrooms due to budget reductions. Study classroom sizes ranged from 16 students to 36 students. Teachers reported working hard to implement the investigations as designed and to provide adequate time for exploration and experimentation, but often felt stretched to achieve "full implementation." For example, one teacher commented that it was very challenging to observe each of the nine groups of four students in her classroom and provide quality feedback to all students, but that ultimately it was “worth the effort to follow the instructional and assessment recommendations, because my students learn so much, and really love science. It is the best part of our day!”

Classroom observation coding. In addition to the qualitative analysis above, we used the observation data to create two quantitative variables to provide another perspective on implementation and to validate log findings. The first variable focused on the *extent to which* teachers implemented ASK/FOSS assessment guidelines and the second characterized the *quality* of the assessment implementation.

The first, which we term fidelity of assessment use, is a summary of whether or not each of assessment components (see Table 20) was in evidence during the observation or follow-up interviews, for example, evidence that a teacher analyzed work in student notebooks, analyzed work on student response sheets, recorded observations of students during class, provided feedback, or engaged students in self assessment. Teachers received a score of “1” for evidence of implementation of the assessment component, a score of “0” if there was no evidence of the teacher using the assessment component. Nine components were used in the analysis, with a possible score range of 0-9 points.

Additionally, a four-point coding scheme was used to rate the *quality* with which each assessment components was used. The maximum possible score for quality of assessment implementation thus was 27 points (9 assessment components x “3” maximum score for each component). Codes were as follows:

Table 21

Classroom Observation: Assessment Use Quality

Score	Description of assessment use
3	Use and analysis of assessment component is detailed and specific.
2	Use and analysis of assessment component is general.
1	Use and analysis of assessment component is broad and unspecified.
0	No use or analysis of assessment component.

Descriptive statistics for the observation variables are displayed in Table 22. Fidelity of implementation scores ranged from 3 to 9, with an average of 7.6 of the assessment components observed in use by teachers. The average observed quality of assessment component use was 16.3, indicating that in general, teachers were using assessment components in ways that were consistent with the FOSS curriculum, that is, in ways that were designed to support student learning. Note that, as might be anticipated, the quality of implementation score shows greater variation than do scores for implementation fidelity.

Table 22

Cohort 1: Observation Descriptive Statistics, Assessment Fidelity and Quality

Variable	*N*	Mean (*sd*)	Total possible	Minimum score	Maximum score
Fidelity of Assessment Implementation: Observation	12	7.6 (2.1)	9	3	9
Quality of Assessment Implementation: Observation	12	16.3 (7.4)	27	3	25

Comparison of treatment and control teachers on both fidelity and quality of implementation favored treatment teachers, but differences were not statistically significant. The lack of significance is not surprising given the small sample size (see Table 23).

Table 23

Cohort 1: Observation, Assessment Fidelity and Quality, Treatment vs. Control

	Control		Treatment		Difference
Variable	*N*	Mean	*N*	Mean	(treatment-control)
Fidelity of Assessment Implementation: Observation total	4	6.75	8	8.00	+1.25
Quality of Assessment Implementation: Observation total	4	14.50	8	17.13	+2.63

Cohort 1 Phone Teacher Interviews

A 50% sample of Cohort 1 treatment and control teachers was selected to participate in phone interviews during the spring of 2010, as they implemented the second module for the year (*Water* or *Structures of Life)*. Interview questions were designed to parallel the classroom observation components and to be carried out in a 30-minute timeframe. Interviews were intended to provide additional data on fidelity of implementation, including teachers thinking and reasoning behind specific instructional and assessment decisions, and in their analysis and interpretation of student work. These processes are often not evident through observations and/or surveys.

Similar to the observation coding, we coded eight FOSS assessment components implemented by both treatment and control teachers (see Table 24), as well as treatment specific assessment components teachers described in interviews. Interviews were first coded according to a "yes/no" scale for use of specific assessments and strategies specified in the FOSS full implementation model. We describe this as *fidelity of implementation*. Next, interviews were coded for the *quality of implementation* of each component as described by the teacher during the interview. These codes were devised by reading the teacher interviews, collecting evidence of the teacher comments, and then coding the interviews. Assessment quality ranged from "0," meaning the assessment or assessment strategy was not used by the teacher, to "3," signifying that the teacher used the assessment component, and provided detailed and specific information about how the tool was used. See Table 24 for details on the assessment component interview quality codes.

Table 24

Assessment Component Quality/Interview

Code	Description and example
3	Use and analysis of assessment component is detailed and specific. E.g., "I recorded observations of students during the investigations, and used these data to help me figure out which students understood the different structures of the crayfish, and the function of each part to provide additional learning experiences for specific students."
2	Use and analysis of assessment component is general. E.g., "I recorded observations of students, and used them to help regroup students."
1	Use and analysis of assessment component is broad and unspecified. E.g., "I made some observations of students but didn't record them in a formal way – kept track in my head."
0	No use or analysis of assessment component. E.g., "No, I didn't make formal observations of students in this module."

Descriptive statistics for the interviews are displayed in Table 25. For interviews, fidelity of implementation scores ranged from 1 (interviewed teacher reported engaging in only one assessment component) to 8 (interviewed teacher reported engaging in all assessment components), with an average of 6.3 assessment components used by teachers. The quality of implementation interview score ranged from 1 to 24; at the low end of the scores, one teacher reported engaging in one assessment component, and that use was described broadly and without specificity. At the other end of the spectrum, another interviewed teacher described use of all eight assessment components with detailed, specific examples of assessment use and analysis. As with the observation ratings, there is greater variability between teachers in ratings of quality than those of implementation fidelity. Note that the data suggest that treatment teachers are implementing the I-Checks and with relatively high quality.

Table 25

Cohort 1: Descriptive Statistics for Interview and Observation Fidelity and Quality

Variable	*N*	Mean (*sd*)	Total possible	Minimum score	Maximum score
All teachers					
Fidelity of Assessment Implementation: Interview	18	6.3 (1.9)	8	1	8
Quality of Assessment Implementation: Interview	18	12.3 (6.2)	24	1	24
Treatment teachers only					
Fidelity of Assessment Implementation: Interview (I-Checks)	10	6.7 (0.5)	7	6	7
Quality of Assessment Implementation: Interview (I-Checks)	10	14.6 (2.7)	21	10	18

Interview results for treatment and control teachers reveal stronger implementation for treatment teachers relative to the control group. Table 26 shows consistently higher scores for treatment teachers on both fidelity of assessment implementation and quality of assessment implementation for both interview and observation measures.

Table 26

Treatment-Comparison Group Implementation: Interview Results

Variable	Control *N*	Control Mean	Treatment *N*	Treatment Mean	Difference (treatment-control)	*T*-value
Fidelity of Assessment Implementation: Interview total	8	5.3	10	7.1	1.8	2.24
Quality of Assessment Implementation: Interview total	8	10.4	10	13.9	3.5	1.22
Fidelity of Assessment Implementation: Observation total	4	6.8	8	8.00	1.2	0.99
Quality of Assessment Implementation: Observation total	4	14.5	8	17.1	2.6	0.56
Fidelity of Assessment Implementation: Observation specific total	4	2.8	8	3.4	0.6	1.25
Quality of Assessment Implementation: Observation specific total	4	6.3	8	7.3	1.00	0.53

Teacher Logs

Teacher Logs were used to measure teachers' use and implementation of the ASK/FOSS curriculum and assessments, and to provide a general gauge of fidelity of implementation of various program constructs and ideas. Refined based on feedback and pilot data from the practice year, the web-based logs asked teacher to report their FOSS/ASK instructional and assessment activities on a weekly basis. The Teacher Logs were available on-line beginning August 18th, 2009. General reporting categories in the teacher log included: (a) amount of teaching time and use of assessment strategies, (b) instructional strategies employed, (c) assessment strategies used, and (d) levels of student understanding.

Log completion rates varied greatly from teacher to teacher, with some teachers completing as few as 2 logs and others completing more than 20. The disparity arises from both variation in the time teachers spent teaching the curriculum unit and from differential compliance with data collection requirements; the difference in number of logs completed introduces variability in the reliability of individual teacher measures. Nonetheless, with teacher as the unit of analysis and mean scores over the course of the unit an indicator of implementation, Table 27 and 28 display descriptive statistics for Teacher Logs from control and treatment teachers in Cohort 1 and Cohort 2 respectively. Across both cohorts, teachers were relatively consistent in the information they reported in the logs from week to week, based on the information and averages they reported for time teaching, assessment strategies used, and time spent looking at student work.

Most teachers reported teaching science about 3 times/week for about 50 minutes. In Cohort 1, it is noteworthy that treatment teachers spent significantly more time looking at student work and were more likely to report spending more than 10 minutes a day in doing so. Other responses were similar across treatment and control responses (see Table 27).

Table 27

2009 – 2010: Cohort 1 Teacher Log Data (All FOSS Modules)

Teacher log questions	Cohort 1 (AZ) Control *N*=14 teachers	Cohort 1 (AZ) Treatment *N*=16 teachers
FOSS time		
Number of times FOSS taught/week	2.9 (0.9)	3.1 (0.8)
Average minutes/day teaching FOSS	49.5 (25)	49.5 (7.8)
Percentage of logs where teachers reported spending more than 40 minutes/day teaching FOSS	76% (0.3)	53% (0.5)
Average minutes/day looking at student work	5.9 (5.8)	10.8 (6.9)
Percentage of logs where teachers reported spending more than 10 minutes/day looking at student work	20% (0.3)	43% (0.4)
Use of assessments*		
Provided feedback to individual student based on analysis of student work	1.2 (0.8)	1.2 (0.7)
Analyzed observations of students	1.6 (1.1)	1.6 (0.8)
Checked on student understandings at the end of an investigation	1.4 (0.9)	0.8 (0.5)
Engaged students in self-assessment of science learning	1.0 (1.1)	0.7 (0.5)
Analyzed student work in science notebooks	1.1 (1.1)	1.4 (0.8)
Analyzed student work on student response sheets	1.3 (0.8)	1.1 (0.8)
ASK (treatment) specific questions		
Used coding guides in the Benchmark Folio to code I-Check items	N/A	0.3 (0.2)
Selected and used a next-step strategy	N/A	0.4 (0.3)
Administered an I-Check Benchmark Assessment	N/A	0.3 (0.2)
Conducted student self-assessment sessions based on I-Check analysis	N/A	0.3 (0.2)

**Note:* scale = number of times/week teacher reported engaging in activities.

Table 28 shows similar patterns among Cohort 2 teachers: treatment teachers (relative to control teachers) also tended to spend more time looking at student work, and a higher proportion reported spending at least 10 minutes doing do. Other aspects of teachers' assessment practices were similar across treatment and control groups.

Teachers' use of I-Check benchmark tools and strategies, provided as part of ASK/FOSS curriculum but not the traditional curriculum, was less frequent than other assessment use, as would be expected. That is, I-Checks were designed to formally check for student understanding at the end of each investigation and investigations typically extend beyond one week. Further, the data suggest that when teachers administered an I-Check, they followed through with using the coding guides, used results to select and use a next-step strategy for instruction and tended to engage students in self assessment sessions based on their I-Check results. It is interesting that I-Check use tends to appear more frequent in Cohort 2 relative to Cohort 1, which may be a function of one state's pacing guidelines for Cohort 2. In that state, for example, Magnetism and Electricity was completed in 6 weeks, versus the more typical 9 to 12 weeks to complete the unit.

Table 28

2009 – 2010: Cohort 2 Teacher Log Data (All FOSS Modules)

Teacher log questions	Cohort 2 (AZ, TX, WA) Control *N*=91 teachers	Cohort 2 (AZ, TX, WA) Treatment *N*=82 teachers
FOSS time		
Number of times FOSS taught/week	3.1 (0.7)	3.4 (0.7)
Average minutes/day teaching FOSS	47.0 (10.35)	50.4 (10.6)
Percentage of logs where teachers reported spending more than 40 minutes/day teaching FOSS	69%	80%
Average minutes/day looking at student work	8.7 (5.1)	12.5 (6.7)
Percentage of logs where teachers reported spending more than 10 minutes/day looking at student work	34%	56%
Use of assessments*		
Provided feedback to individual student based on analysis of student work	1.8 (0.9)	1.8 (0.9)
Analyzed observations of students	2.2 (0.9)	2.3 (1.0)
Checked on student understandings at the end of an investigation	2.0 (0.9)	1.7 (1.0)

Teacher log questions	Cohort 2 (AZ, TX, WA) Control *N*=91 teachers	Cohort 2 (AZ, TX, WA) Treatment *N*=82 teachers
Engaged students in self-assessment of science learning	1.4 (1.1)	1.1 (0.9)
Analyzed student work in science notebooks	1.7 (0.9)	2.1 (0.8)
Analyzed student work on student response sheets	1.6 (0.8)	1.9 (0.8)
ASK (treatment) specific questions		
Used coding guides in the Benchmark Folio to code I-Check items	N/A	0.5 (0.5)
Selected and used a next-step strategy	N/A	1.0 (0.8)
Administered an I-Check Benchmark Assessment	N/A	0.6 (0.5)
Conducted student self-assessment sessions based on I-Check analysis	N/A	0.5 (0.5)

**Note:* scale = number of times/week teacher reported engaging in activities.

To better understand the log data, we conducted a principle component analysis to inform construction of an implementation variable based on log data. The analysis revealed that a three-factor model accounts for most of the variance in the observed data and the factor loading structure, and makes sense theoretically. Table 29 summarizes the results of these analyses. The set of factors is derived from individual items on Questions 2, 3, and 4 (Q2, Q3, and Q4) plus the aggregated items related to teacher weekly involvement in FOSS. We calculated the aggregated items from Q2, Q3, and Q4 by dividing each item with the aggregated item Q1B_SUM. In this way, we were able to understand each teacher's daily involvement in each classroom assessment component.

Table 29

Teacher Logs: Principle Component Analysis (Rotated Factor Pattern: data from all three modules)

Assessment component	Aggregated items (Q_ave=Q_ave/Q 1B_SUM)	Factor 1 "Assessment"	Factor 2 "End of Investigation	Factor 3 "FOSS Intensity"
Used "At a Glance"	Q2A_ave	**0.68**	0.18	0.03
Planned & used assessment	Q3A_ave	**0.69**	0.40	-0.09
Analyzed student work in notebook	Q3b_ave	**0.79**	0.34	0.04
Analyzed student work on response sheets	Q3c_ave	**0.72**	0.52	0.06

Assessment component	Aggregated items (Q_ave=Q_ave/Q 1B_SUM)	Factor 1 “Assessment”	Factor 2 “End of Investigation	Factor 3 “FOSS Intensity”
Analyzed observations of students	Q3d_ave	**0.74**	0.48	-0.09
Recorded and used assessment information on informal data chart	Q3F_ave	**0.65**	0.42	0.23
Provided feedback to individual student based on analysis of student work	Q3g_ave	**0.67**	0.54	0.01
Provided feedback to the entire class based on analysis of student work	Q3I_ave	**0.64**	**0.61**	-0.09
Retaught content	Q3K_ave	**0.61**	0.45	0.13
Checked on students’ ideas at end of Investigation	Q4a_ave	0.40	**0.77**	-0.16
Engaged Ss in self-assessment	Q4b_ave	0.32	**0.86**	0.05
*Administered I-check	Q4c_ave	0.18	**0.89**	0.05
sum (#time FOSS taught/week)	Q1B_SUM	**0.76**	0.21	-0.35
Average (minutes/day on FOSS)	Q1C_AVE	0.01	0.00	**0.91**
Average (minutes/day analyzing student work)	Q1D_AVE	**0.68**	0.05	0.39
Variance explained	72%			

*Treatment teachers only

As shown in Table 29, most items obtain high factor loadings on Factor 1. In particular, items on teachers’ on-going assessment (from Q3 on the log) are concentrated on Factor 1. We label this factor as the “assessment” factor. For Factor 2, items on assessment practices at the end of each investigation (Q4) all obtain high factor loadings. Factor 2 can be labeled as the “end of investigation” factor. Consistent with last year’s principle component analysis, Q1C_ave, is the only item that has high loading on Factor 3. This factor can be regarded as “intensity of engagement in FOSS” factor, which is the average time per day spent on FOSS. As would be expected, the reliability of scores based on these factors is very high: standardized coefficient alphas are .94 and .90 for Factors 1 and 2 respectively (Factor 3 is a one time scale).

Table 30 compares scores based on each of the factors for treatment and control teachers in each cohort by state. The factor scores are normally distributed, with mean equal to 0 and standard deviation equal to 1. For Factor 1, while treatment scores are consistently higher than control scores across all groups, statistically significant differences were found for Cohort 1 teachers, and Cohort 2 teachers in Texas. For Factor 2, treatment scores are consistently lower than control scores for all groups, with statistically significant difference found for Cohort 1 and two of the three Cohort 2 groups. Scores on Factor 3 are more mixed.

Table 30

Implementation Factor Scores for Treatment and Control Teachers

Teacher log data		Factor 1 scores: Assessment			Factor 2 scores: End of investigation			Factor 3 scores: FOSS intensity			*N*	
State	Cohort	Control	Treatment	Difference Treatment/Control	Control	Treatment	Difference	Control	Treatment	Difference	Control	Treatment
AZ	1	-0.82	-0.07	0.75*	0.08	-0.92	-1.00*	-0.02	0.13	0.15	14.00	16.00
AZ	2	-0.44	0.23	0.67	-0.24	-0.28	-0.04	-0.71	0.07	0.78*	10.00	9.00
TX	2	-0.19	0.85	1.04*	0.70	-0.31	-1.01*	-0.32	0.11	-0.43*	62.00	57.00
WA	2	-0.82	-0.38	0.44	0.21	0.71	0.50*	0.75	0.21	-0.54	19.00	16.00

Statistically significant at the alpha <0.05 level.

Validity of Implementation Variable: Relationships Between Measures

We conducted correlation analyses to explore the relationships among the teacher log items, interview items and observation items. The analysis drew on total scores for fidelity and quality variables from the observations and interviews, and scores for each of the factors identified in the logs. In addition, we included aggregated items from the logs characterizing theoretically important aspects of assessment use (i.e., use of feedback and time teachers' spending analyzing student work). Because observations and interviews were conducted during different modules, we correlated scores with logs for the relevant units. That is, all observations were conducted during the Magnetism and Electricity Module and thus we correlated observation scores with log scores from that module.

Table 31 shows the significant correlation coefficients found between observation and log scores for the Magnetism and Electricity Module. Results show moderately strong relationships between log scores on Factor 1 and observations measures of quality of implementation and observations of teachers' use of feedback. Similar correlations were found for assessment components from the log focusing on how much time teachers spend daily outside of class assessing students' work, frequency of analysis of student notebooks, and use of feedback with observation scores for assessment quality with the fidelity and quality of teachers' use of feedback.

Table 31

Correlations between Observation and Log Variable (Magnetism and Electricity)

Item	Time on analysis (Q1D_AVE)	Analysis of student notebooks (q3b_ave)	Feedback (q3g_ave)	Factor1
Assessment Quality: T observation total	0.41	0.62*	0.78*	0.75*
Fidelity of Assessment Feedback: T observation	0.64*	0.57	0.65*	0.82*
Quality of Assessment: use of feedback: T observation	0.38	0.62*	0.73*	0.71*

*Statistically significant at alpha <0.05 level.

In contrast, correlations between interview and log indices for the second ASK/FOSS module, which generally was *Structure of Life*, were not statistically significant. The only exception was the relationship between log data on time on analysis (Q1D_AVE) and the interview total score on fidelity of implementation (r=.49).

We also examined correlations among scores summarized over all modules, which include observations of Magnetism and Electricity, interviews associated with Structure of Life (and one teacher who implemented the Water module), and all log responses. Results shown in Table 32 generally show an absence of relationship between log variables and quality of assessment implementation, as measured by observations and interviews, and moderate relationships between the primary factor emerging from the logs, Factor 1, and interview and observation ratings of fidelity of implementation. Selected items from the log (i.e., minutes teachers spend a day analyzing student work) show similar relationships to the observation and interview implementation ratings. Note that Factor 2 and Factor 3 scores from the logs show no relationship with interview or observation scores.

Table 32

Statistically Significant (<.05) Correlations Based on Teacher Log Factor Scores, Interviews, and Observations

Teacher logs	Interviews and observations	Correlation coefficient
Q1D_AVE (minutes on analysis of student's work)	Fidelity of implementation, interview total score	0.50
Q1D_AVE (minutes on analysis of student's work)	Fidelity of implementation, observation total score	0.59
Q1D_AVE (minutes on analysis of student's work)	Fidelity of implementation, observation specific items	0.68
Q3G_ave Provided feedback to individual students (days/week)	Assessment quality, interview total score	0.49
Factor1 (assessment factor)	Fidelity of implementation, interview total scores	0.55

Correlations between teachers' interview scores and classroom observation scores are generally high, despite the different module contexts for each. Table 32 presents the details. Note in particular the high correlations between the quality ratings from each instrument.

Table 33

Statistically Significant (<.05) Interview and Observation Correlations

	Interview Total, Fidelity of Implementation	Interview Total, Assessment Quality	Interview Total for Treatment-only items, Assessment Quality
Observation Total, Fidelity of implementation	0.93	0.86	0.75
Observation Total, Assessment Quality	0.86	0.93	0.82
Observation Total for Treatment-only items, Assessment Quality	0.70	0.87	0.87
Total Score for Observation specific items, Fidelity of implementation	0.82	0.84	0.94
Total Score for Observation Specific items, Assessment Quality	0.81	0.94	0.92

Finally, we explored the correlation between the teacher content survey and the other measures; note that the correlations between measures of the same content area are high. Table 33 displays the correlations that are statistically correlated. The correlations between measures for the same content areas are high. The correlation matrix for the open-ended items on the teacher content survey and other measures can be found in Appendix C, with significant correlations highlighted.

Implementation of Study Groups

A final implementation component of the embedded assessment system in the ASK/FOSS (treatment teachers only) is the use of Study Groups. These collaborative meetings were designed to help teachers deepen their knowledge and use of a coding (scoring) guide, and by extension, support the development of teachers' content and pedagogical content knowledge. At the study "kick-off" Cohort 1 teachers had the opportunity to review and revisit the Study Group protocol, and observe a mock Study Group presented by professional developers. In preparation for the Study Group, treatment teachers coded student work from an I-Check, recorded scores, and selected items to discuss with other teachers. Cohort 2 teachers were also provided an opportunity to observe a Study Group session, and to practice using the Study Group protocol and tools.

Interviews, observations and Teacher Logs provide modest evidence of Cohort 1 and 2 involvement in Study Groups. On average, teachers met three times in Study Group, spending approximately forty-five minutes collaborating with colleagues, scoring student work, and discussing patterns and trends in the data. Interviews suggest that some teachers found the Study Groups added an important dimension of understanding and skill in instructional and assessment practices. One teacher observed that working with colleagues helped to illuminate conceptual challenges faced by all students, while another teacher found value in exchanging strategies for "next steps" as students engaged in self-reflection and refinement of selected items on the I-Checks.

Conclusion

Overall, Year 3 of the FOSS efficacy study was successful in a number of areas. The year provided an additional opportunity to test and refine our measures, and the results suggest that we have strengthened their technical quality. We have refined the Teacher Logs, and psychometric and theoretical analyses of teacher responses reveal a primary factor that reflects teachers' on-going assessment practices. Because a primary purpose of the new ASK/FOSS curriculum is to strengthen teachers' knowledge and use of assessment, we believe this factor can serve as a general measure of implementation. We have analyzed interview and observations to differentiate two constructs, one representing *fidelity of implementation* – the extent to which teachers implemented ASK/FOSS guidelines with regard to assessment – and a second factor, which represents *quality* of implementation of the assessments. These quantitative indices were used to compare Cohort 1 treatment and control teachers and to cross validate instruments. While it is difficult to find great consistency with the small samples available, we believe these results are very promising. Teacher Logs and other measures were moderately correlated in expected areas, and the relationships were relatively strong between observations and log data on the same module. The strongest relationships were found between interview and observations results, which presents a challenge for next year's data collection, as both sources are relatively cost-intensive compared to the collection of log data.

Cohort 1 teachers reported and were observed implementing FOSS with a relatively high degree of fidelity, and reported a generally positive impact on students learning. Treatment and control teachers described a higher degree of comfort, familiarity and ease with curriculum, assessments, and requirements of the study than in the practice year. Log completion rates improved for this group during year 2 of the study, perhaps due to greater specificity in the logs themselves to describe implementation or because of additional training provided to teachers during the fall "kick-off" session.

Treatment-control comparisons for Cohort 1 teachers, who have completed their study year, showed promising findings for ASK/FOSS. Logs and interview findings showed significant differences favoring the treatment group with regard to assessment implementation. Given the size of the Cohort 1, such results are very encouraging.

Changes in Cohort 1 teacher content survey results are likewise encouraging, indicative of progress in teachers' content, pedagogical and pedagogical content knowledge. Teaching the FOSS Modules for two consecutive years appears to have a positive impact on teacher knowledge. Of particular interest is the impact of assessment use for teachers on their capacity to analyze and interpret student responses, as well as figure out appropriate next instructional steps – treatment teachers showed higher pre/post growth in these areas, a possible consequence of their use of the ASK curriculum and embedded assessment system.

For Cohort 2, the practice year of the study provided successes and challenges as teachers learned how to use the Teacher Logs, and became more familiar with the pace and requirements of participation in the study. Support personnel provided additional resources and guidance to both cohorts, and served an important role in communication and accountability to the study.

Data Collection and Analysis for 2010-2011

Data collection for Cohort 1 teachers and students is now complete. When analyses are complete and student data are available for Cohort 1 classrooms, CRESST will use Hierarchical Linear Modeling to conduct preliminary analyses of treatment effects, the relationship between implementation fidelity and quality and student learning and the factors (e.g., teacher content-pedagogical knowledge, student demographics) that are related to implementation and learning effects.

Data collection for Cohort 2 will be completed using a similar combination of data collection procedures that were used during the 2009–2010 school year for Cohort 1. That is, based on the most recent plan, teachers will be asked to complete electronic logs weekly as well as to complete end of year content surveys; an anticipated 50% selected sample will participate in interviews and a small sample of teachers will be observed. However, based on the available budget, these plans may need to be scaled back.

References

ASK (Assessing Science Knowledge). (2005). *FOSS Magnetism and electricity module.* Nashua, NH: Delta Education. (Developed at the Lawrence Hall of Science, University of California at Berkeley).

ASK (Assessing Science Knowledge). (2005). *FOSS Structures of life module.* Nashua, NH: Delta Education. (Developed at the Lawrence Hall of Science, University of California at Berkeley).

ASK (Assessing Science Knowledge). (2005). *FOSS Water module.* Nashua, NH: Delta Education. (Developed at the Lawrence Hall of Science, University of California at Berkeley).

FOSS Magnetism and electricity module, teacher guide (2005). Nashua, NH: Delta Education. (Developed at the Lawrence Hall of Science, University of California at Berkeley).

FOSS Structures of life module, teacher guide (2005). Nashua, NH: Delta Education. (Developed at the Lawrence Hall of Science, University of California at Berkeley).

FOSS Water module, teacher guide (2005). Nashua, NH: Delta Education. (Developed at the Lawrence Hall of Science, University of California at Berkeley).

Heller, J. I., Daehler, K. R., Shinohara, M., & Kaskowitz, S. R. (2004, April 2). *Fostering pedagogical content knowledge about electric circuits through case-based professional development.* Paper presented at the annual meeting of the National Association for Research in Science Teaching (NARST), Vancouver, Canada.

Olson, J. F., Martin, M. O., & Mullis, I. V. S. (2008). *TIMSS 2007 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, & National Assessment of Educational Progress (NAEP). (2007). *Science Assessment*. Retrieved. from http://nces.ed.gov/nationsreportcard/itmrlsx/default.aspx.

White, R. T., & Tisher, R. P. (1986). Research on Natural Sciences. In *Handbook of research on teaching,* 3rd Edition. M. Wittrock, Editor. Macmillan Publishing Company, NY, NY.

Appendix A:
ASK/FOSS Data Collection Plan

Item	*Control*	*Treatment*	Information Type	Total Sample Size
Case Teachers	3 total schools (1 each in AZ + WA + TX) with a range of 1 or 2 teachers/school. Some schools only have 1 FOSS teacher. Schools will be representative of the sample (demographics, performance, and level of project implementation). **Rationale:** Limited observations of control teachers. Important to provide context for FOSS implementation.	3 schools each from AZ + WA + TX = 9 total schools, range of 1 or 2 teachers/school. Some schools only have 1 ASK teacher. Schools will be representative of the sample (demographics, performance and level of project implementation). **Rationale:** Case approach allows for more in-depth analysis and documentation of teachers' FOI and ASK assessment strategies. Observations provide data on characteristics of implementation and basis for in-depth questions about assessment practices.	Data on fidelity of implementation from selected schools. Provide context for describing the study and examples from selected schools.	12 schools total = 3 control + 9 treatment.
Pre-observation interview	3 Pre-observation interviews with teachers at case schools. **Rationale:** Purpose is to establish context for observation, and understand Ts assessment strategies.	9 pre-observation interviews with teachers at case schools. **Rationale:** Purpose is to establish context for observation, and understand Ts assessment strategies.	Provide background information to observer; establish guidelines for observation.	12 schools = 3 control + 9 treatment.
Classroom observation	Classroom observations at 3 schools.	Classroom observations at 9 schools. **Rationale:** Purpose is to observe implementation of ASK, and ASK assessment strategies in evidence during instruction.	FOI, background and context on implementation of modules and assessments.	12 schools = 3 control + 9 treatment.
Post-observation interview	3 total post-observation interviews.	9 total post-observation interviews. Purpose is to review ASK assessment strategies observed during instruction, and probe teachers' thinking and understanding of the assessments.	Retrospective data on Ts assessment practices during lesson observed.	12 schools = 3 control + 9 treatment.
Study Group Observation/ Interview	N/A	1 Study Group Observation/Case School, 3 in each state. Total of 9 Study Group Observations/Interviews.	Provide context and data on nature of Study Groups. Link to classroom observations where possible.	9 schools.
Sub-sample of Case Study Schools	3 Control Schools.	9 Treatment Schools.		12 schools total.
Phone	30 - 50% of entire sample of control	30 - 50% of entire sample of treatment teachers		50% of schools

Item	*Control*	*Treatment*	Information Type	Total Sample Size
Interview	teachers “reduced schema” = targeted questions about assessment practices based on Teacher Logs.	employing a “reduced schema” = targeted questions about assessment practices based on Teacher Logs.		(minimum of 2 teachers/school) **Note: will review sampling strategy at conclusion of AZ data collection.
Strategy Summary	FOI will be addressed through Teacher Logs and surveys. Reportable data will include Teacher Logs and surveys.	More extensive, in-depth observations through cases to document context and implementation information—variety of data sources to triangulate. Observation data won't be separately reportable in the sense of being representative of all Ts, but will more provide anecdotes, hypotheses, and data to fuel interpretation of more quantitative findings. Data from observations, interviews and Study Groups will be used in conjunction with reportable data from Teacher Logs and surveys.		
Pre/post Teacher Survey	All teachers.	All teachers.	Background information (education, experience), current teaching and assessment knowledge, and practices.	All teachers.
Content Survey	All teachers.	All teachers.	M&E content knowledge, analytic and interpretive skills for understanding student thinking, pedagogical content knowledge.	All teachers.
Teacher Logs	All teachers.	All teachers.	Quantitative info on FOI, typical practices and approaches to implementation.	All teachers.

Appendix B:
Data Collection Instruments

1. Pre-Institute Teacher Survey
2. Teacher Content Knowledge Survey
3. Classroom Assessment Observation Protocol
4. Interview Protocol (4a.ASK, 4b.FOSS)
5. Teacher Logs, Control and Treatment (Magnetism and Electricity, Water, Structures of Life)
6. Teacher Study Group Logs (Treatment only)
7. FOSS Full Implementation Guidelines

Pre-Institute Teacher Survey

June, 2008

Dear ASK/FOSS Participant:

Attached is a teacher survey we'd like you to complete. The survey is one method for us to collect information regarding teachers' ideas and understandings of the FOSS modules and background information on your teaching experiences and ideas. The information you provide us in this survey is critical to our understanding of how and in what ways teachers think about teaching and assessment, and its various uses and functions in classrooms. Your views will also help us to formulate recommendations for the FOSS modules in the future. You may choose not to answer questions or stop taking the survey at any time.

Responses to this questionnaire will be held strictly confidential. Teachers will not be identified by name or school, and only we at WestEd and CRESST/UCLA will have access to completed questionnaires. Please take a few minutes to fill out this survey.

The value of our work depends on the quality of information we receive. We understand the many demands on your time and appreciate the time and energy that thoughtful responses require.

Thank you in advance for your assistance, and congratulations on your successful participation in the ASK/FOSS study.

Sincerely,

Joan Herman Ellen Osmundson Steve Schneider Mike Timms

If you have questions or comments, please contact:
Ellen Osmundson
925-253-1522
eosmundson21@comcast.net

1. Background Information

Please indicate your responses by checking, circling or filling in the blanks.

1.1 Sex:	[ ]M	[ ]F	
1.2 Ethnicity	[ ] White	[ ] Latino/a, Hispanic	[ ] Native American
	[ ] African American	[ ] Asian	[ ] Other___________

Academic/Professional Background

1.3a What is the highest degree you have received?

[ ] Bachelor's + teaching credential
[ ] Bachelor's + credential + units beyond
[ ] Master's
[ ] Master's + units beyond
[ ] Doctorate
[ ] Other (specify)______________

1.3b Indicate the content/subject(s) for each of your degrees:

[ ] Bachelor's (specify)______________
[ ] Doctorate (specify)______________
[ ] Master's (specify)______________
[ ] Other (specify)______________

1.4 Please indicate which teaching credentials you have and specify the content area of specialization. (Check **ALL** that apply.)

[ ] General Elementary
[ ] General Secondary
[ ] Special Emergency
[ ] Multiple Subject
[ ] Single Subjects
If yes, which subject/s:__________________________
[ ] Bilingual
[ ] Administrative
[ ] Other________________________

1.5 How many years have you been teaching elementary school?

____________________ years

1.6 How many years have you been teaching the FOSS program?

____________________ years

1.7 How much does your school leadership support science teaching at your **school**?

Very limited support		Moderate support		Extensive support
1	2	3	4	5

1.8 How much does your district leadership support f science teaching in your **district**?

Very limited support		Moderate support		Extensive support
1	2	3	4	5

Science and Assessment Professional Opportunities

1.9 Have you recently participated or are you currently participating in any of the following organizations? Please check **ALL** that apply.

[ ] NSTA (National Science Teachers Association) [ ] IRA (International Reading Association)

[ ] ASTA (Arizona Science Teachers Association)

[ ] Other: please describe____

____

1.10 In the past 2 years, approximately how many hours have you participated in professional development activities for science teaching (not including the current institute?) ___________________ hours

2. Instructional and Assessment Time

2.1 In using a typical FOSS module, how many times on average per week do you teach science?
___________________ times per week

2.2 In teaching a typical FOSS module, what is the average length of time per lesson?

- ❑ 20 minutes or less
- ❑ 21-40 minutes
- ❑ 41-60 minutes
- ❑ More than 60 minutes

2.3 How many weeks total (excluding breaks) do you spend, from start to finish, on a typical FOSS module?

- ❑ Less than 8 weeks
- ❑ 8 weeks
- ❑ 9 weeks
- ❑ 10 weeks
- ❑ 11 weeks
- ❑ 12 or more weeks

2.4 To what extent do you typically implement ASK/FOSS modules according to the **scope and sequence** outlined in the Teacher's Manual?

Implement FOSS module with major changes to scope and sequence		Implement FOSS modules with modest changes to scope and sequence		Implement FOSS modules with few if any changes to scope and sequence
1	2	3	4	5

2.5 To what extent do you typically implement FOSS **assessments** according to the process outlined in the Teacher's Manaul?

Implement FOSS assessments with major changes to the processes outlined in the Teacher's Manual		Implement FOSS assessments with moderate changes to the processes outlined in the Teacher's Manual		Implement FOSS assessmentswith few if any changes to the processes outlined in the Teacher's Manual
1	2	3	4	5

2.6 How often do you use the following FOSS activities as an indication of how well students understand the concepts:

		Never or hardly ever	Once per module	Every other week	Weekly	Daily	N/A
a)	Teacher observations: general	1	2	3	4	5	N/A
b)	Teacher observation: specific	1	2	3	4	5	N/A
c)	Student sheets (used by students during investigation to organize data)	1	2	3	4	5	N/A
d)	Response sheets (used by students to record observations and write explanations)	1	2	3	4	5	N/A
e)	Performance assessments	1	2	3	4	5	N/A
f)	End-of-module assessments	1	2	3	4	5	N/A

Comments:

3. Assessments Knowledge and Practices

Assessment Resources

3.1 How familiar are you with the following assessment resources?

		Not at all familiar		Some-what		Very familiar	N/A
a)	National Science Education Standards, National Resource Council (NRC)	1	2	3	4	5	N/A
b)	AZ Science Standards and Benchmarks	1	2	3	4	5	N/A
c)	Classroom Assessment and the National Science Education Standards, (book by NRC)	1	2	3	4	5	N/A
d)	Other resources: list and rate familiarity	1	2	3	4	5	N/A

Comments:

3.2 Please rate yourself along the following dimensions:

		Weak		**Moderately strong**		**Very strong**	**N/A**
a)	Knowledge/understanding of grade level science **standards and benchmarks**	1	2	3	4	5	N/A
b)	Confidence in teaching science	1	2	3	4	5	N/A
c)	Knowledge of a wide variety of assessment strategies and techniques	1	2	3	4	5	N/A
d)	Knowledge of how students' ideas develop in science	1	2	3	4	5	N/A

Comments:

3.3 Assessment System

How often do you do the following:

		Hardly ever	Occasionally	Sometimes	Usually	Always	N/A
a)	Set **daily learning** goals for student learning	1	2	3	4	5	6
b)	Set **unit** goals for student learning	1	2	3	4	5	6
c)	Communicate learning goals to students	1	2	3	4	5	6
d)	Make sure your assessments are aligned with your learning goals	1	2	3	4	5	6
e)	Assess students' knowledge prior to introducing a new module	1	2	3	4	5	6
f)	Use multiple assessment methods to gauge learning	1	2	3	4	5	6
g)	Coordinate items on daily, weekly and unit assessments to gauge how student understandings are developing	1	2	3	4	5	6
h)	Use assessment results as a basis for evaluating student progress toward learning goals	1	2	3	4	5	6

Comments:

3.4 FOSS Unit Goals: Self-Assessment

3.4a How well do you understand the learning goals for the FOSS module on Magnetism and Electricity?

No clear understanding of the learning goals of the FOSS module Magnetism and Electricity (i.e., don't know the big ideas or major concepts)		Moderately good understanding of the goals of the FOSS module Magnetism and Electricity (i.e., have some idea of the big ideas and major concepts)		Very good understanding of the goals of the FOSS module Magnetism and Electricity (i.e., fully understand the big ideas and major concepts)	N/A
1	2	3	4	5	6

Comments:

3.4b How well do you understand the learning goals for the FOSS module on Water?

No clear understanding of the learning goals of the FOSS module “Water” (i.e., don’t know the big ideas or major concepts)		Moderately good understanding of the goals of the FOSS module “Water” (i.e., have some idea of the big ideas and major concepts)		Very good understanding of the goals of the FOSS module “Water” (i.e., fully understand the big ideas and major concepts)	N/A
1	2	3	4	5	6

Comments:

3.4c How well do you understand the learning goals for the FOSS module on Structures of Life?

No clear understanding of the learning goals of the FOSS module “Structures of Life” (i.e., don’t know the big ideas or major concepts)		Moderately good understanding of the goals of the FOSS module “Structures of Life” (i.e., have some idea of the big ideas and major concepts)		Very good understanding of the goals of the FOSS module “Structures of Life” (i.e., fully understand the big ideas and major concepts)	N/A
1	2	3	4	5	6

Comments:

3.5

When teaching a FOSS module, to what extent do you:

		Hardly ever	Occasionally	Sometimes	Usually	Always	N/A
a)	Listen/ask questions as students work to gauge their understandings	1	2	3	4	5	6
b)	Expect all students to learn the concepts and ideas	1	2	3	4	5	6
c)	Analyze and interpret students’ ideas based on a developmental framework of how science understandings develop	1	2	3	4	5	6
d)	Analyze and interpret **whole group discussions** for general patterns of understanding	1	2	3	4	5	6
e)	Analyze and interpret **small group discussions** and work for specific student understandings	1	2	3	4	5	6
f)	Analyze and interpret individual work and responses for student understandings	1	2	3	4	5	6
g)	Use assessments in ways that allow	1	2	3	4	5	6

	all students to “show what they know”						
h)	Other: specify	1	2	3	4	5	6

Comments:

3.6 Please indicate how you are currently using assessment in science.

When you teach a FOSS module, to what extent are you currently ***usingresults from your FOSSassessments to****:*

		Minimal extent		Moderate extent		Great extent	N/A
a)	Understand students’ science ideas from your assessment evidence	1	2	3	4	5	6
b)	Provide **group feedback** to students	1	2	3	4	5	6
c)	Provide **individual feedback** to students	1	2	3	4	5	6
d)	Plan next steps in instruction	1	2	3	4	5	6
e)	Involve students in analyzing and interpreting their work	1	2	3	4	5	6
f)	Use assessments in ways that allows all students to “show what they know”	1	2	3	4	5	6
g)	Other: specify	1	2	3	4	5	6

Comments:

Assessment Tool Quality

3.7 Please indicate your perceptions of the assessment tools you generally use to evaluate student learning in science.

To what extent are your tools:

		Very limited extent		Moderate extent		Great extent	N/A
a)	Based on scientifically accurate information	1	2	3	4	5	6
b)	Aligned with the science concepts in the module	1	2	3	4	5	6
c)	Reliable and accurate						
d)	Provide you the information you need to plan next steps in your instruction	1	2	3	4	5	6
c)	Provide the information you need to help individual students clarify their understandings	1	2	3	4	5	6
d)	Valid for the reason you are using them (does it test/measure what you thought it would?)	1	2	3	4	5	6
e)	Fair (allows all students to show what they know in a different ways)	1	2	3	4	5	6
f)	Designed to accommodate learners	1	2	3	4	5	6

with various needs

Comments:

Science Content Knowledge

3.8 How well qualified do you feel to teach 4th grade students about the following topics?

		Not at all qualified		Somewhat qualified		Very qualified	N/A
a)	Magnetism and electricity	1	2	3	4	5	6
b)	Magnetic forces	1	2	3	4	5	6
c)	How electrical ciruits are designed	1	2	3	4	5	6
d)	How electricity in circuits can produce magnetic effects	1	2	3	4	5	6
e)	Water	1	2	3	4	5	6
f)	The water cycle	1	2	3	4	5	6
g)	Properties of water	1	2	3	4	5	6
h)	The nature of Earth materials: what are they and their properties	1	2	3	4	5	6
i)	How water interacts with the Earth's crust, oceans, and atmosphere	1	2	3	4	5	6
j)	Structures of Life	1	2	3	4	5	6
k)	Seeds: what they are, what they look like and how they grow	1	2	3	4	5	6
l)	The life cycle of a plant	1	2	3	4	5	6
m)	Animal habitats	1	2	3	4	5	6
n)	How an organism's structures help it survive in its habit	1	2	3	4	5	6

Comments:

3.9 How and in what ways do you provide feedback to your students about their learning based on FOSS assessments?

3.10 How, if at all, do you use assessment results to plan next steps in your instruction?

4.0 Student Outcomes

4.1 To what extent are the FOSS assessments aligned with your Local (District/State) science assessments?

Not aligned		Somewhat aligned		To a great extent
1	2	3	4	5

Please explain:

4.2 To what extent do students' inquiry skills typically improve as a result of using the FOSS modules?

Not at all		To a moderate extent		To a great extent
1	2	3	4	5

Please explain:

4.3 To what extent do your students typically engage in self-assessment as part of FOSS assessments?

Not at all		To a moderate extent		To a great extent
1	2	3	4	5

Please explain:

4.4 How much time do you typically spend analyzing and reflecting on students' responses to a FOSS assessment students complete during an investigation activity?

- ❑ less than 10 minutes
- ❑ 10 minutes
- ❑ 20 minutes
- ❑ 30 minutes
- ❑ 40 minutes
- ❑ 50 minutes
- ❑ 60 minutes or more

4.5 Approximately how much time do you spend analyzing and interpreting students' responses on the FOSS **end-of section** or **end-of-investigation** assessment tasks?

- ❑ less than 30 minutes
- ❑ 30 minutes
- ❑ 60 minutes
- ❑ 90 minutes
- ❑ More than 90 minutes

4.6 To what extent does scoring and analyzing FOSS assessments provide you with information about student learning and progress?

Not at all		To a moderate extent		To a great extent
1	2	3	4	5

Please explain:

4.7 To what extent does scoring and analyzing FOSS assessments provide information you use to guide subsequent instruction (next instructional steps)?

Not at all		To a moderate extent		To a great extent
1	2	3	4	5

Please explain:

5. Summary Questions

In this section of the survey, please answer the questions as completely as possible in the space provided.

5.1 What are your current views about the role of assessment in student science learning?

____

5.2 What are your goals for participation in this project?

Thank you for completing the survey.

TEACHER CONTENT SURVEY

FOSS STUDY

PURPOSE OF THIS INSTRUMENT

This measure is designed to collect information about teacher understandings of magnetism and electricity and approaches teachers use to understand student thinking. Results from the survey will help us to better understand how FOSS works to help students learn science.

INSTRUCTIONS

1. You have been alloted 30–45 minutes to complete this measure. However, if you wish, you may use more time during your break in order to finish it. You may choose to not answer questions and/or stop your work at any point during the time period.

The content survey includes questions with a wide range of difficulty, and we expect you to encounter items for which you may not know the answers. If you are not sure of an answer, ***please make your best guess***—there is no penalty for guessing.

2. Please fill in your name and ID numbers below and your ID on the next page.

First name Last name Date

Your ID Number: T ☐ ☐–☐ ☐–☐ ☐ ☐

IMPORTANT:

To keep your data confidential, this cover sheet with your name will be removed upon receipt by the research staff, leaving only your ID number on the next page of the survey. This cover sheet will be stored in a locked cabinet, separate from the completed surveys.

Your ID Number: T ☐ ☐–☐ ☐–☐ ☐ ☐

1.11 Julie placed a paper clip, piece of cardboard, and magnet together like you see in the picture.

Why did the paper clip stay in place next to the cardboard instead of falling to the floor? Choose the **best** answer.

❍ A. The paper clip is made of iron and so is the magnet.

❍ B. The magnetic field goes around the cardboard and makes the paper clip stay there.

❍ C. The magnet has a magnetic field that is not blocked by the cardboard.

❍ D. The electric force field makes the paper clip attract to the magnet.

1.12 Arthur was playing with magnets. He had one magnet on the table, and one in his hand. As he moved the magnet in his hand closer to the one on the table, the magnets suddenly snapped together.

a. Explain why the magnets snapped together even though they were not touching.

Here are two students' responses to question 1.12:

Student 1 Response: Both magnets are made of iron, and the magnets are both facing south and south.

Student 2 Response: The magnets snapped together because the electric fields got close.

b. What inferences can you draw about the students' understanding of magnetism and electricity? What do these students know? What do these students not know/need to learn?

c. If these students were in your class, what would you do next in your instruction to help the students learning progress?

1.21 A nail that was stuck to a permanent magnet picked up a small metal washer. The nail could pick up the metal piece because:

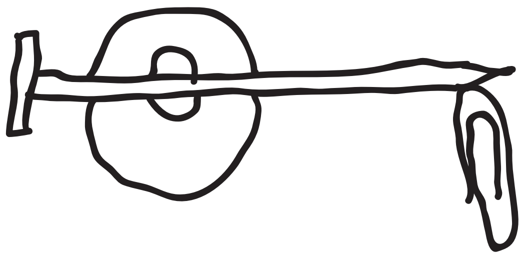
❍ A. Nails have magnetic fields.

❍ B. Magnetism was induced in the nail.

❍ C. The nail and the washer are both made of iron.

❍ D. The washer is still in the range of the magnetism.

1.22 Anne is investigating objects and magnets. She made this observation in her science journal.



"I was surprised! A nail was stuck to the magnet. When I accidentally touched the nail to a paper clip, the paper clip stuck to the nail. I wonder why that happened?"

a. Explain to Anne why the paper clip stuck to the nail. Use diagrams or pictures if necessary.

Anne and her friend were asked by her teacher why they thought the paper clip stuck to the nail. Here are their responses to the question:

Anne's response: The paper clip turned into a magnet too.

Anne's friend's response: The nail gets stuck on the magnet, and the nail turns into a magnet, so the paper clip can stick on the nail.

b. What inferences can you draw about the students' understanding of magnetism and electricity? What do these students know? What do these students not know/need to learn?

c. If these students were in your class, what would you do next in your instruction to help the students learning progress?

1.31 **a.** Complete the following table. Put an "X" in the second column of the table if the object sticks to a magnet. Put an "X" in the third column of the table if the object conducts electricity.

Object	Sticks to a magnet	Conducts electricity
Iron nail		
Plastic straw		
Steel wire screen		
Wooden craft stick		
Brass ring		
Rubber band		
Copper penny		
Piece of aluminum foil		

b. Why did you choose the objects that you did in the "Sticks to a magnet" column? Use diagrams or pictures to show your thinking.

c. Why did you choose the objects that you did in the "Conducts electricity" column? Use diagrams or pictures to show your thinking.

1.32 Here is how one student completed the table.

Object	Sticks to a magnet	Conducts electricity
Iron nail	X	X
Plastic straw		
Steel wire screen	X	X
Wooden craft stick		
Brass ring	X	X
Rubber band		
Copper penny	X	X
Piece of aluminum foil	X	X

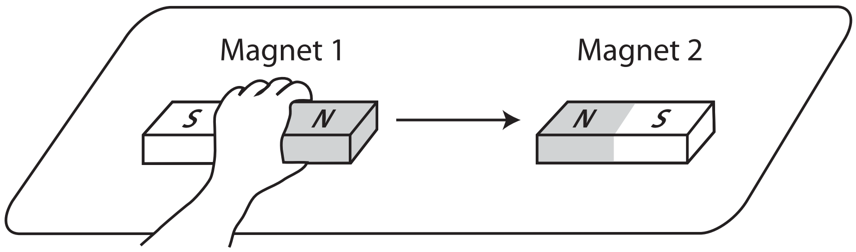
Here are one student's responses to questions 1.31b and 1.31c (see page 5):

Student 1 Response:
1.31 b. These things stick to the magnet because they are all metal.
1.31 c. These things are all made of metal and metal conducts electricity.

a. What inferences can you draw about the students' understanding of magnetism and electricity? What do these students know? What do these students not know/need to learn?

b. If these students were in your class, what would you do next in your instruction to help the students learning progress?

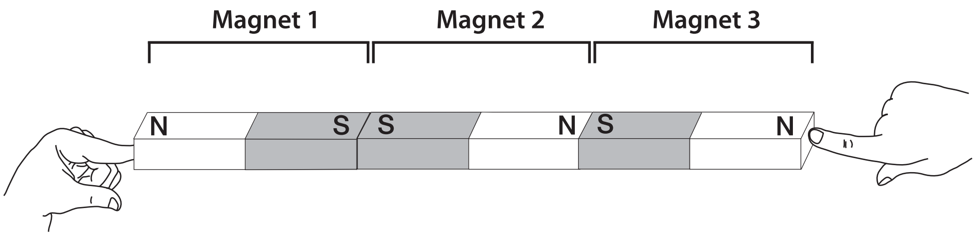
1.41 The picture below shows Maria pushing Magnet 1 toward Magnet 2 on a smooth table. Both magnets are lying on a smooth table.



a. What will happen as Magnet 1 moves towards Magnet 2?

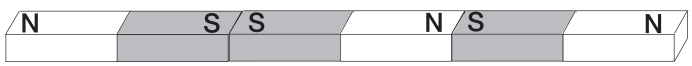
b. Why will this happen?

1.42 Three bar magnets are held together as shown in the picture below.

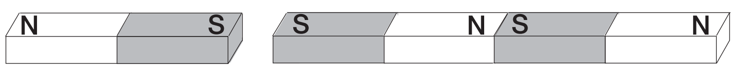


a. What will the magnets do when they are released? Circle the correct answer.

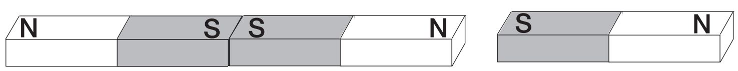
A.



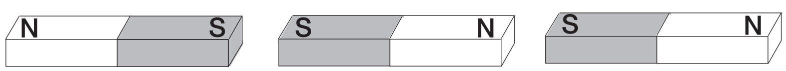
B.



C.

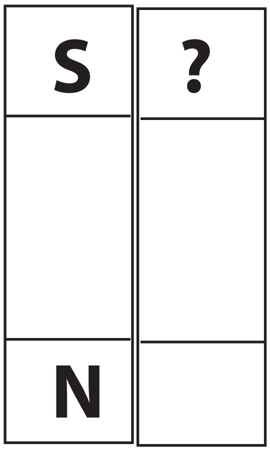


D.



b. Why does that happen?

1.46 Lisa found a magnet with no labels on the poles. She found another magnet with correctly labeled poles and put the magnets together. They **attracted**.



a. The pole labeled with the "?" is most likely which pole?

❍ A. south pole
❍ B. north pole
❍ C. not enough information provided

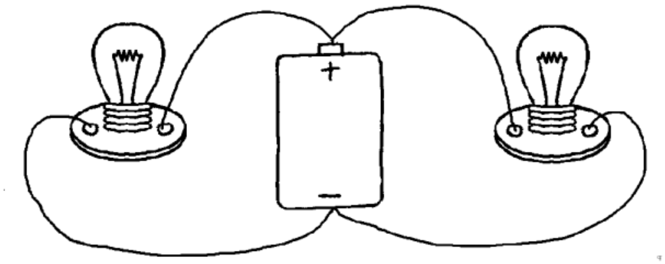
b. Why? Please explain your answer.

WRAPPING IT UP

1. What is/are the key concept/s addressed by the assessments in Section 1?

2. Why is it important for students to learn these magnetism and electricity concepts?

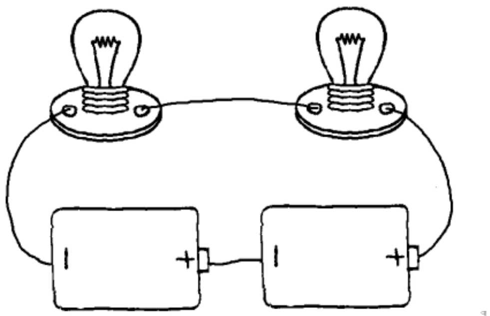
2.11 Look at the picture below. What kind of circuit is this?



How do you know?

❍ A. network circuit
❍ B. series circuit
❍ C. parallel circuit
❍ D. short circuit

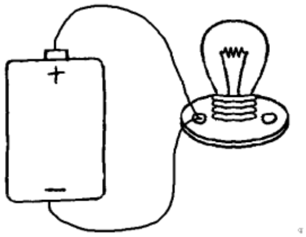
2.12 Look at the picture below. What kind of circuit is this?



How do you know?

❍ A. simple circuit
❍ B. series circuit
❍ C. parallel circuit
❍ D. short circuit

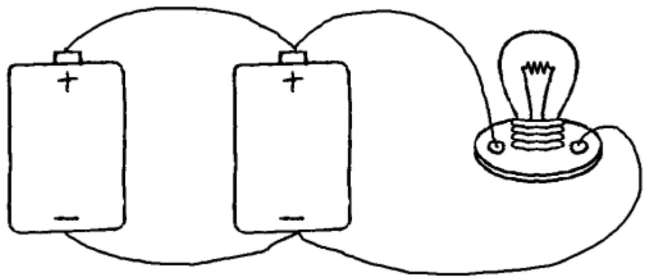
2.13 Look at the picture below. What kind of circuit is this?



How do you know?

❍ A. simple circuit
❍ B. series circuit
❍ C. parallel circuit
❍ D. short circuit

2.14 Look at the picture below. What kind of circuit is this?



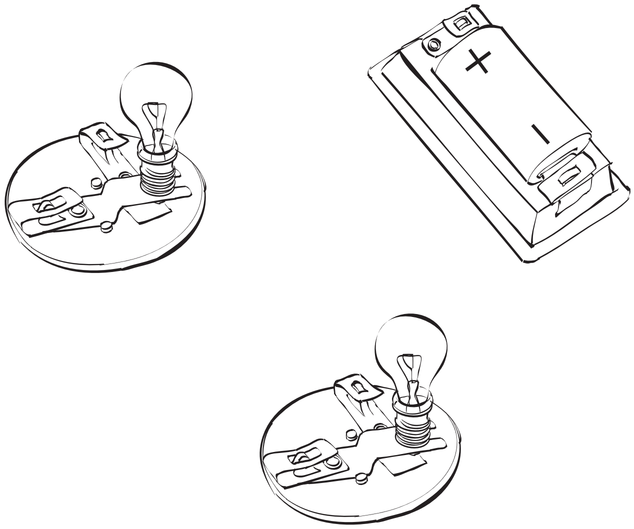
❍ A. simple circuit

❍ B. network circuit

❍ C. series circuit

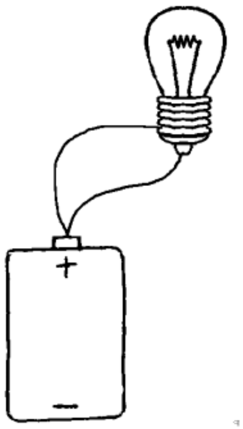
❍ D. parallel circuit

2.15 a. Draw in lines representing wires to make a ***parallel*** circuit.



Explain your drawing: what features make this a parallel circuit?

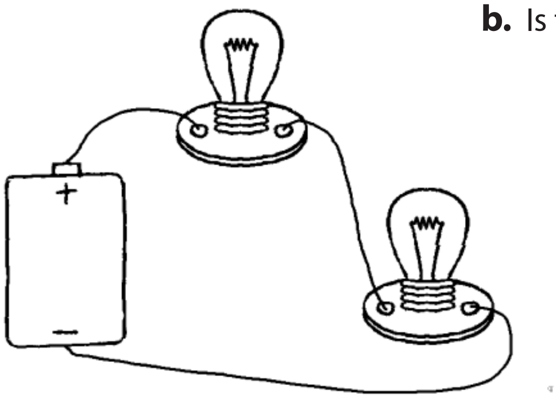
2.21 Look at the picture below.



a. Will the bulb light? ❍ Yes ❍ No

b. Is the circuit complete? ❍ Yes ❍ No

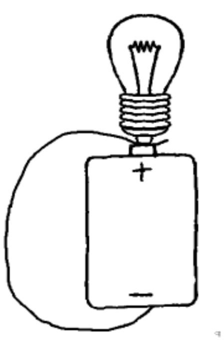
2.22 Look at the picture below.



a. Will the bulbs light? ❍ Yes ❍ No

b. Is the circuit complete? ❍ Yes ❍ No

2.23 Look at the picture below.



a. Will the bulb light? ❍ Yes ❍ No

b. Is the circuit complete? ❍ Yes ❍ No

c. Explain why you think the circuit is or is not complete.

2.24 Look at the picture below. The round object in the middle of the picture is an empty bulb holder.

a. Will the bulb light? ❍ Yes ❍ No

b. Explain why you think the bulb will or will not light.

This is how a Student 1 responded to question 2.24.

a. Will the bulb light? ❍ Yes ⊗ No

b. Bulb won't light because it's not connected to the battery.

This is how a Student 2 responded to question 2.24.

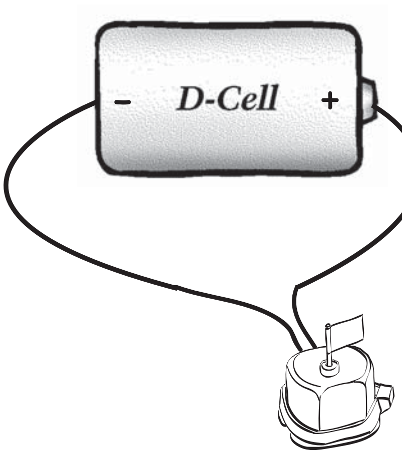
a. Will the bulb light? ❍ Yes ⊗ No

b. Bulb won't light because it's a short circuit.

c. What inferences can you draw about the students' understanding of magnetism and electricity? What do these students know? What do these students not know/need to learn?

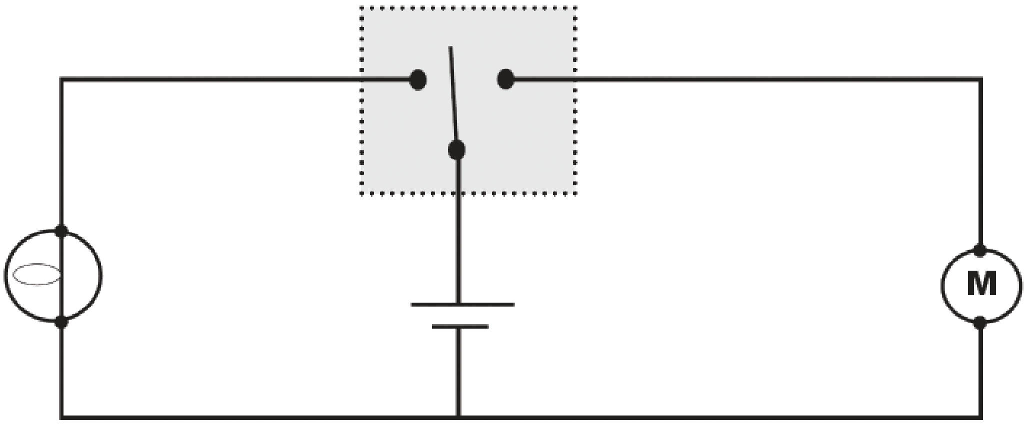
d. If these students were in your class, what would you do next in your instruction to help the students learning progress?

2.31 Draw arrows on the picture to show which direction electricity will flow through the circuit to run the motor.



Explain your answer.

2.32 Denise wants to build a circuit that will light up a bulb and run a motor at the same time. She drew the diagram of the circuit she planned to build. She used a special switch in the circuit. The switch is shown in the gray box.



a. Look at the diagram Denise drew. Explain to her why you think her circuit would or would not work the way she wants it to work.

2.33 Below are two student's responses to question 2.32.

Student 1 response: I think it would work because all the parts of connected. But it might not work because the battery might not have enough juice to carry all on one circuit.

Student 2 response: It probably won't because the energy can't go two different ways.

a. What inferences can you draw about the students' understanding of magnetism and electricity? What do these students know? What do these students not know/need to learn?

b. If these students were in your class, what would you do next in your instruction to help the students learning progress?

2.41 Electricity can be changed into other forms of energy. Complete the sentences below:

a. The bulb in a lamp changes electric energy into ______________________________

____

b. A motor changes electric energy into ______________________________

____

2.42 Which of the following items converts electric energy into motion?

❍ A. light switch
❍ B. electric stove
❍ C. light bulb
❍ D. electric fan

2.43 When an electric stove is turned on, most of the incoming electrical energy changes into:

❍ A. heat energy
❍ B. light energy
❍ C. mechanical energy
❍ D. sound energy

2.44 Which of the following items converts electric energy into light?

❍ A. light switch
❍ B. doorbell
❍ C. light bulb
❍ D. electric fan

2.45 When an electric fan is running, most of the incoming electric energy is converted into:

❍ A. heat energy
❍ B. light energy
❍ C. motion energy
❍ D. sound energy

2.46 Household appliances convert electricity into one or more different forms of energy. An electric fan can best be described as converting electricity into:

❍ A. heat energy only
❍ B. heat energy, and sound energy
❍ C. heat energy, sound energy and motion energy
❍ D. heat energy, sound energy, motion energy and chemical energy

WRAPPING IT UP

1. What is/are the key concept/s addressed by the questions in Section 2?

2. Why do students need to know these concepts about magnetism and electricity?

3.11 Annie had three rivets. One was copper, one was iron and one was steel. Which rivet or rivets could she use to make an electromagnet? Why?

3.12 Here are two students' response to question 3.11

Student 1: Annie should use the iron and steel rivets because they conduct electricity and they stick to magnets.

Student 2: Annie could use the iron, copper or steel rivets because they are all metal.

a. What inferences can you draw about the students' understanding of magnetism and electricity? What do these students know? What do these students not know/need to learn?

b. If these students were in your class, what would you do next in your instruction to help the students learning progress?

3.21 Imagine you have the following materials: a large iron nail, several permanent magnets, lots of insulated wire, a D-cell and a switch.

a. Describe one way to make the nail a temporary magnet.

b. Describe another way to make a temporary magnet.

3.31 Samuel Morse, the inventor of the telegraph, had a problem. His telegraph's signal was too weak. He needed a stronger electromagnet. What are two ways he might have used to increase the strength of the electromagnet for his telegraph?

3.41 Wendy is making an electromagnet. First, she wrapped a long, insulated wire around an iron nail. What should Wendy do to complete the electromagnet?

Here are two student responses to question 3.41:

Student 1: Attach the wire to the D-cell and switch, rub the magnet on the nail a few times and then try it.

Student 2: Wendy should connect the iron nail to the D-cell to make a complete circuit.

a. What inferences can you draw about the students' understanding of magnetism and electricity? What do these students know? What do these students not know/need to learn?

b. If these students were in your class, what would you do next in your instruction to help the students learning progress?

3.42 Which of the following materials is **NOT** necessary to build an electromagnet?

❍ A. a magnet

❍ B. a steel rivet

❍ C. a D-cell battery

❍ D. wire

WRAPPING IT UP

1. What is/are the key concept/s addressed by the assessments in Section 3?

2. Why do students need to know these concepts about magnetism and electricity?

ASK/FOSS Classroom Assessment Observation Protocol

Observation Notes

	Descriptions
Observer	
Date	
School	
Teacher	
Module & Investigation	
Concepts Addressed	
Lesson	
Introduction: Focus question	
Activity 1 *Describe the activity, what teacher is doing, what students are doing, interactions.*	
Activity 1 Assessment *To what extent is T involved with assessing Ss?* (1=not at all, 3=moderate extent,	

5=great extent)	
Activity 2 *Describe the activity, what teacher is doing, what students are doing, interactions.*	
Activity 2 Assessment *To what extent is T involved with assessing Ss?* (1=not at all, 3=moderate extent, 5=great extent)	
Activity 3 *Describe the activity, what teacher is doing, what students are doing, interactions.*	
Activity 3 Assessment *To what extent is T involved with assessing Ss***?** (1=not at all, 3=moderate extent, 5=great extent)	
Other observational data	

(fill out as observing)	
Classroom description	
Assessment materials *in evidence (per activity/task if appropriate)*	
Other: please indicate	

Note: observer should take notes during the observation and complete the scaled items at the conclusion of the lesson and/or after reviewing notes.

	Check if observed	Description	Congruence/Alignment with FOSS/ASK assessment system **1=not at all** **3=moderate** **5=to a great extent** **see below**
Prior to the lesson/investigation			
Used the "At a Glance" to review science content and assessment opportunities for teaching and assessment			
During the lesson/investigation			
Analysis and Interpretation			
Analyzed students' science notebooks			
Used a scoring guide to analyze response sheets			
Recorded observations of students' during class			
Analyzed **student work** for patterns and trends			
Analyzed **observations** for patterns and trends			
Feedback to Students			

	Check if observed	Description	Congruence/Alignment with FOSS/ASK assessment system 1=not at all 3=moderate 5=to a great extent see below
Individual Individual students provided ongoing, clear feedback regarding progress toward targeted goals.			
Small Group Targeted, specific, descriptive feedback is provided to students working in small groups regarding progress towards targeted goals.			
Whole Class Targeted, specific, descriptive feedback is provided to whole class regarding progress towards targeted goals.			
Notebooks Provided feedback to students on notebook entries			
Provided opportunities for students to work in small groups to discuss ideas			
Asked open-ended questions			
Guide for Instruction			
Planned and implemented additional instruction based on **observations of students** during class			
Planned and implemented additional instruction based on **assessment results**			

	Check if observed	Description	Congruence/Alignment with FOSS/ASK assessment system **1=not at all** **3=moderate** **5=to a great extent** **see below**
End of lesson/investigation			
Checked on students' understandings of science concepts			
Engaged students in self-assessment of science learning			
Other: please specify			

ASK Study Research: Phone Interview

[Purpose: provide more detailed and specific information on teacher assessment practices, based on guidelines from LHS on "full implementation" model, and certain components of CRESST's Quality Assessment Model]

Note: it may be helpful to provide the teacher with a copy of the interview protocol to help him/her follow the questions and the conversation.

Introduction

(Interviewer introduces self)

Hi. As you know, we're conducting interviews with teachers in the ASK/FOSS study to help us better understand your use of FOSS and how are you are assessing students. This interview to bring me up to date on the _______ (Module and Investigation).

Do I have your permission to audiotape this conversation? I will use the tape only to ensure I have complete notes. As we outlined in our information letter and permission documents, your confidentiality is assured, and you have the right not to answer any questions and to terminate the conversation at any time.

Do you have any questions before we begin?

Interview Questions

1. **General Update:** I want to get a general sense of the ______ (Module and Investigation) you've been teaching. Note: Section 1 should take 2 - 3 minutes.
   a. Based on the information you've provided in the Teacher Logs, I see that you have just finished ______ Module and Investigation (interviewer needs to check Teacher Logs in advance of interview). How are things going – what has worked well so far with this Module and Investigation? What has been a challenge? (keep very brief)
   b. Which assessments have you used to date (check all that apply: pre-test, I-checks, student response sheets, notebooks). In general, how are things going with assessing students' learning – what has worked well so far? What has been a challenge? (keep very brief)

2. <u>Use of Assessments</u>

Now I'd like to ask you more specific questions about your ASK assessment practices.

Note: interviewer will take brief notes here to describe the process. Audiotape can be used to supplement the details, but does not have to be transcribed verbatim.

In this current Investigation, have you:	***Yes/No***	**If yes, then:** **a. *how* did you use the tool/do it?** **b. what did you find out about student learning from this process or work?** **c. what do the results mean for your teaching?**
a. Analyzed work in students' science notebooks		
b. Analyzed student work on the response sheets		
c. Recorded observations of students' during class		
d. Analyzed **student work** for patterns and trends		
e. Analyzed **observations** for patterns and trends		
f. Planned and used a next-step strategy based on student work		
g. Provided feedback to students about their work and learning		

3. **<u>End of Investigation</u>**

Note: these questions apply (need to be asked) only if the teacher indicates that s/he is at the conclusion of an Investigation.

Interviewer: Next, I'd like to ask you about the end of investigation assessments. (refer back to information in #1 to guide next set of questions).

In this current Investigation, have you:	***Yes/No***	**If yes,** **a. *how* did you use the tool/do it?** **b. what did you find out about student learning from this work?** **c. what do the results mean for your teaching?**
a. Administered the I-Check Benchmark Assessment		
b. Used coding guides in the Benchmark Folio to analyze I-Check		
c. Recorded I-Check data on the Benchmark Coding sheets		
d. Conducted student self-assessment session after I-Checks were returned to students		
e. Checked student reflections (revisions) after self-assessment session		
f. Made instructional decisions based on I-Check results		
g. Other: please specify		

4. **Study Groups**

In this current Investigation, have you:	***Yes/No***	**If yes,** **a. describe what you did** **Note: see other specific questions below**
a. met as a Study Group		
b. scored work in your Study Group		**a. describe what you did** **b. what did you find out about student learning from this work?** **c. what do the results mean for your teaching?**
c. figured out next steps strategies based on the combined student work		**a. describe what you did**
d. planned next instructional steps		**a. describe what you did**
e. other: please describe		**a. describe what you did**

5. **Wrap Up**

Do you have any other questions or comments to add? Thanks very much for your time.

FOSS Study Research: Phone Interview

[Purpose: provide more detailed and specific information on teacher assessment practices, based on guidelines from LHS on "full implementation" model, and certain components of CRESST's Quality Assessment Model]

Introduction

(Interviewer introduces self)

Hi. As you know, we're conducting interviews with teachers in the ASK/FOSS study to help us better understand your use of FOSS and how are you are assessing students. This interview to bring me up to date on the _______ (Module and Investigation).

Do I have your permission to audiotape this conversation? I will use the tape only to ensure I have complete notes. As we outlined in our information letter and permission documents, your confidentiality is assured, and you have the right not to answer any questions and to terminate the conversation at any time.

Do you have any questions before we begin?

Interview Questions

6. **General Update:** I want to get a general sense of the ______ (Module and Investigation) you've been teaching. Note: Section 1 should take 2 - 3 minutes.
   a. Based on the information you've provided in the Teacher Logs, I see that you have just finished ______ Module and Investigation. How are things going – what has worked well so far with ? What has been a challenge? (keep very brief)
   b. Which assessments have you used to date (check all that apply: pre-test, student response sheets, notebooks, other). In general, how are things going with assessing students' learning – what has worked well so far? What has been a challenge? (keep very brief

7. Use of Assessments

Now I'd like to ask you more specific questions about your assessment practices when teaching FOSS.

Note: interviewer will take brief notes here to describe the process. Audiotape can be used to supplement the details, but does not have to be transcribed verbatim.

In this current Investigation, have you:	***Yes/No***	**If yes,** **a. *how* did you use the tool/do it?** **b. what did you find out about student learning from this work?** **c. what do the results mean for your teaching?**
a. Analyzed students' science notebooks (if applicable)		
c. Used a scoring guide (or coding guide) to analyze response sheets		
e. Recorded observations of student's during class (e.g., in small groups, 1:1 conversations)		
f. Analyzed **student work** for patterns and trends		
e. Analyzed **observations** for patterns and trends		
f. Planned further instruction based on patterns and trends in student work (specify which work)		
g. Provided feedback to students about their work and learning		

8. <u>End of Investigation</u>

Note: these questions apply (need to be asked) only if the teacher indicates that s/he is at the conclusion of an Investigation.

Interviewer: Next, I'd like to ask you about the end of investigation assessments. (refer back to information in #1 to guide next set of questions).

In this current Investigation, have you:	***Yes/No***	**If yes,** **a. *how* did you use the tool/do it?** **b. what did you find out about student learning from this work?** **c. what do the results mean for your teaching?**
a. Checked on students' understandings at the end of a lesson or an investigation (describe)		
b. Engaged students in self-assessment of science learning		
c. Other: please specify		

9. <u>Wrap Up</u>

Do you have any other questions or comments to add? Thanks very much for your time.

WestEd FOSS Teacher Study

About FOSS | Documents | **Teacher Logs** | My Profile

Log Completed · Log Not Completed · Log Past Due · Current Log

Available Logs: Week of 09/07/2009 Week of 09/14/2009

The teacher log below is intended for the Water module.

If you are NOT CURRENTLY teaching the **Water** module, please update your profile to reflect your current module.Your respons
to these questions will be confidential except for two items, which are clearly marked in red below.

Teacher Log for the Week of 09/07/2009

☐ I did not teach FOSS this week

General

1a) Which Investigation/s did you work on this week?
(check the appropriate boxes)

Water:

☐ Survey (pretest)	☐ 3.2
☐ 1.1	☐ 3.3
☐ 1.2	☐ 3.4
☐ 1.3	☐ 4.1
☐ 2.1	☐ 4.2
☐ 2.2	☐ 4.3
☐ 2.3	☐ 4.4
☐ 3.1	☐ Posttest

(This information will be made available to support staff and contractors to allow them to better support you.)

1b) On which days did you teach FOSS this week? ☐ Mon ☐ Tue ☐ Wed ☐ Thu ☐ Fri

1c) On the days that you taught science, approximately how many minutes did you spend teaching FOSS? (record number of minutes in each box)

Mon	Tue	Wed	Thu	Fri

1d) This week, approximately how many minutes each day did you spend looking at student work after teaching FOSS? (record number of minutes in each box)

Mon	Tue	Wed	Thu	Fri	Wkend

This week, during FOSS instruction, how many days did you engage in the following activities?

Resources **DAYS USED THIS WEEK**

2)	Used the "**At a Glance**" to review focus (investigation) questions, science content, and assessment opportunities for your teaching	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
Assessments		**DAYS USED THIS WEEK**
3a)	Planned and used an **assessment** for the lesson (e.g., observation, response sheet, student sheet, performance assessment)	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
3b)	Analyzed student work in **science notebooks**	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
3c)	Analyzed student work on the **response sheets**	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
3d)	Analyzed **observations** of students	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
3e)	Recorded and used assessment information on an **informal data chart**	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
3f)	Provided **feedback to individual students** based on analysis of student work	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
3g)	Provided **feedback to the entire class** based on analysis of student work	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
3h)	**Retaught** content based on analysis and interpretation of student work	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
3i)	What did you learn about students' understanding of science concepts from your analysis of student work? Please provide examples and specific details.	
End of Investigation		**DAYS USED THIS WEEK**
4a)	Checked on student understandings at the end of an Investigation	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
4b)	Engaged students in self-assessment of science learning	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
4c)	Checked students' reflections after self-assessment	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
4d)	What did you learn about students' understanding about science concepts at the end of the investigation? Please provide examples and specific details.	
Comments		
5a)	What percentage of your students do you think understand the core concepts of the Investigation(s) you taught this week?	
5b)	Do you have any questions or feedback about your experience with the project this week? This question is not confidential and responses will be made available to support staff and contractors to allow them to better	

support you.

Once you submit this log you will not be able to go back and edit it. Please make sure all your answers are entered correctly!

Submit Teacher Log

FOSS Teacher Study

About FOSS | Discussion | Documents | **Teacher Logs** | My Profile

Log Completed   Log Not Completed   Log Past Due   Current Log

Available Logs: Week of 09/07/2009   Week of 09/14/2009

The teacher log below is intended for the Water module.

If you are NOT CURRENTLY teaching the **Water** module, please update your profile to reflect your current module.Your respons to these questions will be confidential except for two items, which are clearly marked in red below.

Teacher Log for the Week of 09/07/2009

☐ I did not teach FOSS/ASK this week

General

1a) Which Investigation/s did you work on this week? (check the appropriate boxes)

Water:

☐ Survey (pretest)
☐ 1.1
☐ 1.2
☐ 1.3
☐ 2.1
☐ 2.2
☐ 2.3
☐ 3.1
☐ 3.2
☐ 3.3
☐ 3.4
☐ 4.1
☐ 4.2
☐ 4.3
☐ 4.4
☐ Posttest

(This information will be made available to support staff and contractors to allow them to better support you.)

1b) On which days did you teach FOSS/ASK this week? ☐ Mon ☐ Tue ☐ Wed ☐ Thu ☐ Fri

1c) On the days that you taught science, approximately how many minutes did you spend teaching FOSS/ASK? (record number of minutes in each box)

Mon	Tue	Wed	Thu	Fri

1d) This week, approximately how many minutes each day did you spend looking at student work after teaching FOSS/ASK? (record number of minutes in each box)

Mon	Tue	Wed	Thu	Fri	Wkend

This week, during FOSS/ASK instruction, how many days did you engage in the following activities?

Resources — **DAYS USED THIS WEEK**

2a) Used the "**At a Glance**" to review focus (investigation) questions, science content, and assessment opportunities for your teaching	N/A 1 2 3 4 5
2b) Used the "**ASK Teacher Guide Insert Pages**" to guide instruction and formative assessment practices	N/A 1 2 3 4 5
Assessments	**DAYS USED THIS WEEK**
3a) Planned and used an **embedded assessment** for the lesson (e.g., notebook sheet, notebook entry, response sheet, observation)	N/A 1 2 3 4 5
3b) Analyzed student work in **science notebooks**	N/A 1 2 3 4 5
3c) Analyzed student work on the **response sheets**	N/A 1 2 3 4 5
3d) Analyzed **observations** of students	N/A 1 2 3 4 5
3e) Analyzed student work for **patterns and trends**	N/A 1 2 3 4 5
3f) Recorded and used assessment information on an **informal data chart**	N/A 1 2 3 4 5
3g) Provided **feedback to individual students** based on analysis of student work	N/A 1 2 3 4 5
3h) Used sticky notes, conferences, etc. to provide **individual feedback** to students based on analysis of student work	N/A 1 2 3 4 5
3i) Provided **feedback to the entire class** based on analysis of student work	N/A 1 2 3 4 5
3j) Selected and used a **next-step strategy**	N/A 1 2 3 4 5
3k) **Retaught** content based on analysis and interpretation of student work	N/A 1 2 3 4 5
3l) What did you learn about students' understanding of science concepts from your analysis of student work? Please provide examples and specific details.	
Benchmark Assessments for Investigations	**DAYS USED THIS WEEK**
4a) Checked on student understandings at the end of an Investigation	N/A 1 2 3 4 5
4b) Engaged students in self-assessment of science learning	N/A 1 2 3 4 5
4c) Administered an I-Check Benchmark Assessment	N/A 1 2 3 4 5
4d) Used **coding guides** in the Benchmark Folio to code I-Check items	N/A 1 2 3 4 5
4e) **Recorded** I-Check codes on the "Summary Coding Sheets"	N/A 1 2 3 4 5

4f)	Conducted **student self-assessment sessions** based on I-Check analysis	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
4g)	Checked students' reflections after self-assessment	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
4h)	Used a **next-step strategy** based on self-assessment sessions	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
4i)	Describe the self-assessment activities you provided for the class.	
4j)	What did you learn about students' understanding about science concepts based on information from the I-Checks? Please provide examples and specific details.	

Comments

5a)	What percentage of your students do you think understand the core concepts of the Investigation(s) you taught this week?	
5b)	Do you have any questions or feedback about your experience with the project this week? This question is not confidential and responses will be made available to support staff and contractors to allow them to better support you.	

FOSS/ASK Study Group

6)	Did you meet with your FOSS/ASK Study Group this week?	○ YES ○ NO

Once you submit this log you will not be able to go back and edit it. Please make sure all your answers are entered correctly!

Submit Teacher Log

WestEd FOSS Teacher Study

Logged in as: **Dan Psomas** [logout]

About FOSS | Discussion | Documents | User Admin | User Tracking | Download | Report Status

FOSS/ASK Study Groups

FOSS/ASK Study Groups 08/31/2009	
An important part of this study involves meeting with a Study Group at least three times during each module to discuss student work. Please answer the following questions about your Study Group work.	
General	
1a) Which module is your Study Group meeting to discuss?	
1b) Which Study Group meeting number is this?	
1c) Approximately how many minutes did the meeting last?	
Please provide the following information about your Study Group: (a) the amount of time you spent on the activities listed below, and (b) a brief description of the process you used	
Choosing Items to Moderate	
2a) Approximate time spent (minutes)	
2b) Brief description of the process	
Coding (Scoring) Student Work	
3a) Approximate time spent (minutes)	
3b) Brief description of the process	
Discussing Student Responses to Questions	
4a) Approximate time spent (minutes)	
4b) Brief description of the process	
Discussing Patterns and Trends in Student Responses	
5a) Approx time spent (minutes)	
5b) Brief description of the process	
Deepening/Increasing Your Content Understanding	
6a) Approx time spent (minutes)	
6b) Brief description of the process	
Planning Next Instructional Steps Based on Discussion of Student Responses	
7a) Approximate time spent (minutes)	

7b)	Brief description of the process	
Other, Please Describe:		
8a)	Approx time spent (minutes)	
8b)	Brief description of the process	
For the section below: 1 = Not beneficial at all; 3= Moderately Beneficial; 5 = Extremely beneficial		
How beneficial was this Study Group session in helping you to:		
9a)	Understand your students' learning	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
9b)	Identify patterns and trends in student responses	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
9c)	Deepen/increase your content understanding	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
9d)	Plan next instructional steps	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
9e)	Learn to use the coding (scoring) guides	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
9f)	Overall benefit of the session	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
9g)	Other (please specify)	○ N/A ○ 1 ○ 2 ○ 3 ○ 4 ○ 5
Summary		
10a)	What, if any, new ideas, strategies, or approaches that were discussed in your Study Group will you implement in **assessing student** work in FOSS/ASK?	
10b)	What are the 3 most important things you learned from the Study Group today?	

Once you submit this log you will not be able to go back and edit it. Please make sure all your answers are entered correctly!

Submit Study Group Questions

Full Implementation of FOSS/ASK Embedded Assessment System

Instruction

- In a full implementation of the ASK curriculum, teachers follow the Teacher Guide including ASK insert pages. They guide the investigations as described in the Teacher Guide, ask all the relevant questions posed in the Teacher Guide, follow the wrap up, read the stories, add concepts and questions as appropriate to the Content/Inquiry chart, and make use of the Word Bank. Teachers use extensions and next-step strategies as resources for additional activities or instructional strategies to support student learning.

Embedded Assessments

- Teachers use the "ASK Teacher Guide Insert Pages" to guide instruction and formative assessment practices.

- Teachers plan an embedded assessment for each lesson (investigation part). Students may answer a focus question, look for patterns from observations/data on a notebook sheet, complete a response sheet, or whatever teachers deems important. Suggestions are made in each *Getting Ready* section in the "ASK Teacher Guide Insert Pages." Embedded assessments are written/drawn by students in their science notebooks.

- Teachers collect the science notebooks at the end of each lesson and spend at least 10 minutes reviewing student work and 5 minutes more reflecting and deciding on next steps. If there is not time to review all notebooks after every lesson, a sampling process can be used to select specific notebooks. Teachers make every effort to look at EVERY students' response when reviewing Response Sheets (more formal embedded assessments that occur once each investigation). Teachers take notes, and develop an informal data chart that displays diagnostic information for each embedded assessment.

Benchmark Assessments

- Teachers administer and code the Survey (pretest) before instruction begins. They use the information gained about prior knowledge to inform instruction, but do not give any codes, scores, or feedback to students. Teachers submit original student work to project staff and maintain copies for use by students at the end of the module.

- Teachers administer I-Checks after each investigation. They code I-Checks outside of class, but make no marks on the students' papers. Codes are kept on a separate spreadsheet. Teachers also make decisions about which items need to be brought back to the students for the self-assessment session. Not all items should be presented to students for reflection, only a sample.

- Teachers initiate a self-assessment session with the students after each I-Check. Items that are used in the self-assessment session are those teachers noted (when coding) that students were having problems answering. Students can also participate in deciding which items to discuss further.

- Essentially, it takes one class period to administer each I-Check and it takes another class period to complete the self-assessment. Questions on the I-Check may be read aloud to students, otherwise they should receive no other help. Notebooks may NOT be used when students take the I-Check, they CAN be used during the self-assessment session.

- After the student self-assessment session, teachers decide whether they need to spend more time on a particular concept students don't understand, but in general, they move on to the next investigation. At a minimum, teachers provide students another opportunity to learn the concepts in the context of the class/group interaction during the self-assessment session.

- After completing the module, teachers administer the Posttest and submit original student work to project staff. Teachers may use the Pretest and the Posttest for students to compare their responses and assess their growth in scientific understanding.

Next-Step Strategies

Next-step strategies are what teachers do...

(1) after looking at each embedded assessment to continue instruction. A next-step strategy is used to help students clarify alternative conceptions teachers may have noticed when reading the students' notebooks, and reviewing work from the lesson;
(2) after the I-Checks. Self-assessment activities are the next-step strategy suggested to follow each I-Check. Teachers may also decide to use other next-step strategies if they find students need more help than the self-assessment provides.

Study Groups

- Teachers meet as a Study Group after each I-Check during the course of a module to discuss student work with team members. Teachers bring copies of student work from a mutually agreed upon I-Check, along with a spreadsheet of codes representing student responses. The group follows a mutually agreed upon protocol to discuss student work and completes a study group log at the end of the session.

Teacher Logs

- *Teachers complete an electronic log on a weekly basis. These logs provide implementation information about teachers' instructional and assessment practices.*

Professional Development for FOSS/ASK Participation

Teachers participate in the following professional development for each of the two project FOSS modules.

- One full day training is provided on module content and refining strategies for management of materials and student activities.
- One full day of training is provided on the assessment system.

Full Implementation of FOSS System

Instruction

- In a full implementation of the FOSS curriculum, teachers follow the Teacher Guide. They guide the investigations as described in the Teacher Guide, ask all the relevant questions posed in the Teacher Guide, follow the wrap up, read the stories, add concepts and questions as appropriate to the Content/Inquiry chart, and make use of the Word Bank. It is not necessary for teachers to use all of the extensions, although these resources serve as potential sources of information for additional activities to support student learning.

Assessment

- Teachers administer the Pretest (Survey), submit the originals to project staff, and may maintain copies for his or her records. (These should not be given back to students until after they have taken the Posttest.)
- Teachers use the response sheets and the student sheets as appropriate to support student learning.
- Teachers administer the Posttest, submit the original to project staff, and maintain copies for students.

Teacher Logs

- *Teachers complete an electronic log on a weekly basis. These logs provide implementation information about teachers' instructional and assessment practices.*

Professional development for FOSS Study Participation

- Teachers participate in one full-day professional development for each of the two project FOSS modules. This professional development focuses on the module content as well as refining the management of materials and the module activities.

Appendix C:
Correlation Matrix - All Instruments

	pre_factor1	pre_factor2	pre_CONTENTmag	pre_CONTENTelec	pre_CONTENTelec2	pre_CONTENTelectrom	post_factor1	post_factor2	post_CONTENTmag	post_CONTENTelec	post_CONTENTelec2	post_CONTENTelectrom
pre_Mag	0.11	0.13	-0.10	-0.09	-0.03	0.21	0.12	0.18	-0.04	0.11	0.25	-0.35
	0.56	0.47	0.54	0.58	0.84	0.18	0.48	0.30	0.84	0.56	0.19	0.06
	31.00	31.00	40.00	40.00	40.00	40.00	35.00	35.00	28.00	28.00	28.00	28.00
pre_Elec	-0.07	0.09	0.05	0.22	0.22	**0.61**	0.01	0.13	0.19	0.34	**0.49**	-0.07
	0.69	0.63	0.77	0.18	0.16	**<.0001**	0.97	0.47	0.34	0.07	**0.01**	0.74
	31.00	31.00	40.00	40.00	40.00	**40.00**	35.00	35.00	28.00	28.00	**28.00**	28.00
pre_ElectroMag	0.00	0.00	0.25	0.04	0.13	**0.70**	0.01	0.10	-0.09	0.31	0.29	**-0.40**
	1.00	0.98	0.12	0.79	0.44	**<.0001**	0.97	0.58	0.64	0.11	0.14	**0.04**
	31.00	31.00	40.00	40.00	40.00	**40.00**	35.00	35.00	28.00	28.00	28.00	**28.00**
post_Mag	-0.01	0.02	**0.46**	0.24	0.20	0.11	-0.05	0.10	**0.55**	**0.60**	**0.61**	0.17
	0.97	0.94	**0.01**	0.21	0.30	0.57	0.79	0.56	**0.00**	**0.00**	**0.00**	0.38
	30.00	30.00	**30.00**	30.00	30.00	30.00	34.00	34.00	**28.00**	**28.00**	**28.00**	28.00
post_Elec	**-0.36**	0.22	**0.37**	0.12	0.06	0.26	-0.26	0.29	0.21	**0.53**	**0.45**	-0.14
	0.05	0.25	**0.04**	0.52	0.77	0.17	0.13	0.10	0.29	**0.00**	**0.02**	0.47
	30.00	30.00	**30.00**	30.00	30.00	30.00	34.00	34.00	28.00	**28.00**	**28.00**	28.00
post_ElectroMag	-0.02	0.29	0.31	0.33	0.22	0.00	-0.10	0.33	**0.46**	**0.53**	**0.60**	**0.40**
	0.92	0.12	0.09	0.08	0.24	0.98	0.57	0.06	**0.01**	**0.00**	**0.00**	**0.04**
	30.00	30.00	30.00	30.00	30.00	30.00	34.00	34.00	**28.00**	**28.00**	**28.00**	**28.00**
pre_content	-0.09	-0.19	0.18	-0.06	-0.03	**0.53**	-0.14	-0.13	0.10	0.15	0.22	-0.30
	0.64	0.31	0.27	0.70	0.86	**0.00**	0.43	0.44	0.60	0.46	0.26	0.12
	31.00	31.00	40.00	40.00	40.00	**40.00**	35.00	35.00	28.00	28.00	28.00	28.00
pre_Analysis	0.11	0.17	-0.06	0.03	0.09	**0.57**	0.16	0.19	0.02	0.25	0.31	-0.28
	0.55	0.37	0.72	0.86	0.57	**0.00**	0.37	0.28	0.91	0.21	0.11	0.14
	31.00	31.00	40.00	40.00	40.00	**40.00**	35.00	35.00	28.00	28.00	28.00	28.00
pre_nextstep	-0.04	0.15	0.13	0.18	0.23	**0.58**	0.03	0.23	-0.01	0.36	**0.51**	-0.29
	0.83	0.42	0.42	0.26	0.16	**0.00**	0.87	0.18	0.95	0.06	**0.01**	0.14

	pre_factor1	pre_factor2	pre_CONTENTmag	pre_CONTENTelec	pre_CONTENTelec2	pre_CONTENTelectrom	post_factor1	post_factor2	post_CONTENTmag	post_CONTENTelec	post_CONTENTelec2	post_CONTENTelectrom
	31.00	31.00	40.00	40.00	40.00	**40.00**	35.00	35.00	28.00	28.00	**28.00**	28.00
post_content	-0.05	0.13	0.20	0.16	0.05	0.02	-0.06	0.16	**0.61**	**0.51**	**0.41**	**0.37**
	0.78	0.49	0.28	0.38	0.80	0.93	0.73	0.37	**0.00**	**0.01**	**0.03**	**0.05**
	30.00	30.00	30.00	30.00	30.00	30.00	34.00	34.00	**28.00**	**28.00**	**28.00**	**28.00**
post_Analysis	-0.18	0.27	**0.41**	0.30	0.24	0.16	-0.20	0.33	**0.50**	**0.66**	**0.64**	0.19
	0.35	0.15	**0.03**	0.11	0.21	0.40	0.25	0.06	**0.01**	**0.00**	**0.00**	0.34
	30.00	30.00	**30.00**	30.00	30.00	30.00	34.00	34.00	**28.00**	**28.00**	**28.00**	28.00
post_nextstep	-0.10	0.09	**0.53**	0.24	0.17	0.15	-0.09	0.18	**0.38**	**0.58**	**0.63**	0.08
	0.60	0.63	**0.00**	0.20	0.37	0.42	0.60	0.30	**0.05**	**0.00**	**0.00**	0.69
	30.00	30.00	**30.00**	30.00	30.00	30.00	34.00	34.00	**28.00**	**28.00**	**28.00**	28.00

Note: Prob > |r| under H0: Rho=0.