Engineering is Elementary:

An Evaluation of Year 6 Field Testing

Dr. Cathy P. Lachapelle
Dr. Christine M. Cunningham
Bharat Battu
Kris Keenan
Engineering is Elementary
The National Center for Technological Literacy
Museum of Science, Boston
December, 2009

Abstract

During the 2008-2009 school year (EiE Year 6), two new Engineering is Elementary (EiE) units were field tested: *Designing Parachutes* and *Designing Solar Ovens*. Pre- and post-assessments were collected from students in field test and control classrooms in five states. In field test classrooms, students participated in an EiE unit and a related science, while in control classrooms, students participated in related science instruction only. Though the data collection and instruments were flawed, results for the evaluation are positive and promising. Compared to control students, EiE students performed significantly better on questions assessing engineering, science, and technology. Both girls and boys improved significantly on the post-assessments after participating in an EiE unit, as did students from all ethnic groups, and students receiving free and reduced-price lunch. Analysis of the Year 6 data has led to numerous improvements in assessment design that will be implemented in EiE Project Year 7.

Introduction

Engineering is Elementary: A Curriculum Development Project

Engineering is Elementary (EiE) is a research-based curriculum development project focused on creating engineering and technology curriculum units to supplement core science instruction. Each EiE curriculum unit is designed to build on and reinforce a science topic through the exploration and development of a related technology. For example, in the *Seeing Animal Sound* unit, students design representations of bird sounds using the engineering design process (EDP). EiE units are created with a four-lesson structure. Lesson one introduces a particular engineering field through an engineering story involving a fictional child character who faces an engineering problem. Lesson 2 uses hands-on activities to explore the field of engineering more broadly. In the third lesson, students use controlled experiments to explore different materials, processes, or design elements. These experiments will inform their final designs. In the fourth lesson, students use the engineering design process to plan, create, test, evaluate and improve their designs[1].

In this report, we analyze results of evaluation of two units developed during 2007-2008 (EiE Year 5) and field tested in 2008-2009 (Year 6):

- *Designing Parachutes*
- *Designing Solar Ovens*

The EiE Research and Evaluation Program

EiE's research and evaluation program has four major goals: 1) to learn about what students (and teachers) across the nation know about and believe about engineering, technology, and the engineering design process; 2) to improve the EiE curriculum by observing lessons in action and collecting feedback from teachers; 3) to develop assessment tools for EiE, school districts, and others to use in evaluating the implementation of EiE in settings large and small; and 4) to evaluate the impact of the Engineering is Elementary curriculum on students' understanding of engineering, technology, and related science topics.

Over the course of 2-3 years, each EiE unit undergoes a rigorous cycle of design, testing, and redesign. During the first year of this cycle, each new unit is drafted and "pilot tested" in several Massachusetts classrooms. Based on the first year of testing, each new unit is revised. In the second year of the design cycle, twelve teachers from each of five states receive copies of each revised unit and assessments for field testing. These field test teachers provide the EiE staff with feedback on the unit and student pre- and post-assessment data and receive from EiE the materials they need to implement the unit in their classrooms; based on the results of field testing each unit is revised again (and if necessary, undergoes limited testing of revisions for a third year).

The first year of EiE unit development (the "pilot" year) is devoted to developing and intensive in-classroom testing of new units by EiE staff with the help of pilot teachers. During this year, the EiE research and evaluation program has two tasks: developing assessments for the new units under development, and formative assessment of those same units. As part of the formative evaluation process, members of the EiE staff observe units in classrooms as teachers teach the units for the first time. The EiE staff use their observations as well as teacher and student feedback to revise the units and assessments.

The second year of EiE unit development (the "field test" year) is devoted to testing the new units in a wider range of classrooms. In each of five states (California, Colorado, Florida, Massachusetts, and

Minnesota), twelve teachers for each unit are found to implement and test the units in their classrooms. These field test teachers attend professional development workshops in their home states where they receive hands-on training in how to implement the units. They collect pre- and post-assessments from their students as well as student demographics; these are sent to EiE where they are scanned and analyzed. Field test teachers also return information about their background, teaching practices, science teaching dates, and feedback on the EiE unit(s) they have taught. Results from the student pre-/post-assessments, as well as teacher feedback, are described in this report.

Methodology

In 2008-2009, two EiE units were field tested: *Designing Parachutes* and *Designing Solar Ovens*.

Assessment Instrument Development

Development of field test instruments for assessing students in 2008-2009 was a two step process. It began in 2007-2008, with the development of pilot versions of the assessments. EiE staff brainstormed a number of questions pertaining to the learning objectives of each unit, as well as related science. From these, sixteen multiple-choice questions were chosen for the pilot assessment for *Designing Solar Ovens*. Twenty-three multiple-choice questions and 11 true/false questions were chosen for the pilot assessment for *Designing Parachutes*. Each pilot assessment was distributed to approximately 80 pilot classroom students (students in classrooms using pilot versions of the EiE units) in grades 3-5 for testing. Based upon student answers to the questions, questions that were too easy (greater than 95% correct) were removed. Other questions were removed because they were no longer relevant to the newest version of the EiE unit; i.e. based on the changes from pilot to field test version of the unit.

The field test version of the *Designing Solar Ovens* unit assessment included 9 multiple-choice questions. The field test version of the *Designing Parachutes* unit assessment included 6 multiple-choice questions and 4 true/false questions.

Data Collection Methodology

In the 2008-2009 school year, two EiE units were tested nationally: *Designing Parachutes*, and *Designing Solar Ovens*. Field testing took place in five states: California, Colorado, Florida, Massachusetts, and Minnesota. This report details the results of pre- and post-assessments of units in field test classrooms and control classrooms. Field test classrooms were classrooms where students were taught one of these two EiE units as well as related science content, while in control classrooms, students were taught related science content (using teachers' standard science curriculum) but not the EiE unit.

Field test teachers were recruited by local field site leaders for EiE in each state. They attended local workshops introducing each new EiE field test unit. They were provided with the materials needed for implementing an EiE field test unit (teacher guide, a classroom pack of storybooks, and a materials kit), and compensated with a small stipend after returning post-assessments and unit feedback.

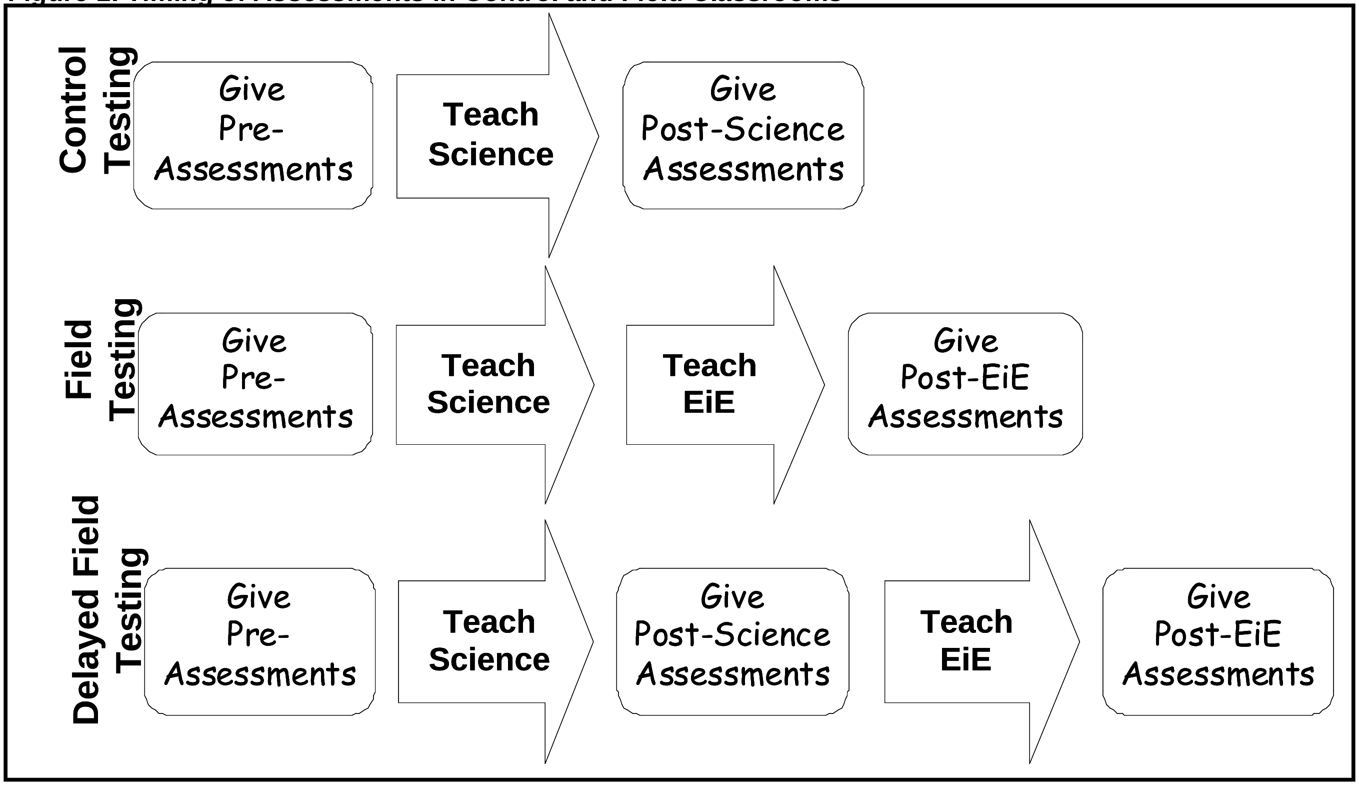
Control teachers were recruited with the promise of admission to a local EiE workshop, EiE field test materials, and stipend. They were compensated with a small stipend after returning post-assessments, and subsequently invited to attend an EiE workshop and receive materials for implementing an EiE field test unit. In this way, we hoped to attract similar populations of teachers for both test and control.

In a given classroom, each student received the same assessments. Each student who participated in the field test program as a field test student or a control student completed one general engineering assessment and/or an “engineering attitudes” survey. The results of the general engineering assessment and the engineering attitudes survey are not included in this report. In addition, each student completed a unit assessment that was related to their science topic and for field test students, the EiE unit they had studied. The unit specific assessments included 9 or 10 science and engineering questions related to the unit.

Whenever possible, students completed pre-assessments in October or November and post-assessments four to eight weeks after completion of the EiE and related science units, in the same academic year. There are, however, varying circumstances for different teachers, so the dates that pre-assessments were given to students vary. For example, some students are taught by science specialists who only see the students for a portion of the year. Some teachers first heard about the project and signed up as field test teachers in the spring of the academic year. No matter when teachers administered their assessments, they were instructed to administer pre-assessments to the students before the students were exposed to the EiE unit or any related science topic. In addition, depending on if a classroom was a control or a test classroom, students needed to complete their post-assessments after they learned about the science topic or completed EiE lessons.

In some classrooms (approximately 10% of the *Designing Parachutes* sample, and 20% of the *Designing Solar Ovens* sample) students were tested *both* as control and test participants. In such classrooms, students were taught the related science unit first; pre-assessments were given before, and post-assessments after this unit—the same as any other control classroom. However, in these classrooms, students then proceeded to take part in EiE unit instruction. Post-assessments were again given to students after the EiE unit was completed. For purposes of distinguishing between the different control conditions, such classrooms were termed “Delayed Field Test” classrooms.

Figure 1. Timing of Assessments in Control and Field Classrooms



The timing of instruction and assessments in control, field test, and delayed field test classrooms is described in Figure 1. EiE students were tested at least twice—once before beginning the science curriculum and/or related Engineering is Elementary unit, and once after instruction was completed—allowing for a test-retest analysis. On the unit-specific assessments, student responses were scored as "correct" or "incorrect" before beginning analysis.

Once all completed assessments from test and control classroom teachers participating in EiE's research for Project Year 6 were collected, they were digitized using an assessment scanner and then imported into a Microsoft Access database. This data was then exported to PASW Statistics version 17.0, together with student demographic data, for analysis. Student responses to each question were recoded as correct/incorrect, and a preliminary scale item consisting of the means of all questions was also computed for each assessment. Data collected from all classrooms was checked for any transcription errors and unusual or missing data by inspecting the means from this all-item scale, for the pre-assessment and post-assessment separately.

Methods of Analysis

Responses from students participating in EiE field tests (called EiE or test below) were compared to responses from a control sample. Both the test sample and the control sample received science instruction after completing the pre-assessments and before completing post-assessments. The test sample completed the EiE curriculum in addition to their regular science curriculum.

Assessment items were combined into scales to test for reliability; principal components factor analysis was run to search for item groupings. No item sub-groups were found for either the *Designing Parachutes* or *Designing Solar Ovens* assessments; the all-item scale for the *Designing Solar Ovens* assessment was rejected because of low reliability. The *Designing Parachutes* all-item scale was analyzed using ANCOVA as most suitable for pre-/post-assessment analysis[2].

Individual items on both assessments were also analyzed using chi-square procedures for dichotomous data. This served as the primary means of analysis for the *Designing Solar Ovens* assessment, and as a comparison to the ANCOVA analysis of the *Designing Parachutes* assessment. Items were analyzed individually for within-group (pre vs. post) differences using McNemar's Test of Symmetry, a cross-tabulation analysis designed for binomial nominal data. Differences between the test and control groups were analyzed using the phi coefficient. This χ-square variant is designed for analyzing dichotomous data; its value approaches that of Pearson's χ-square for high values of N, an expectation which was confirmed in this analysis.

Results

The National Field Test Sample: 2008-2009

1590 EiE (test) students and 548 control students completed the *Designing Parachutes* assessment; 767 EiE students and 506 control students completed the *Designing Solar Ovens* assessment. The number of control students completing each assessment was slightly more than half the number of test students. The largest numbers of assessments were collected from Florida and Massachusetts. Grade 5 students are more common in the student samples than other grades. Sample sizes for each assessment by state and grade are summarized in Table 1 and Table 2 below.

Table 1. Sample Size by Assessment, by State

		CA	CO	FL	MA	MN	Total
Designing Parachutes	Control	52	167	254	75	-	548
	Test	84	176	464	583	283	1590
Designing Solar Ovens	Control	109	81	-	135	181	506
	Test	84	216	-	301	166	767

Table 2. Sample Size by Assessment, by Grade

		Grade 3	Grade 4	Grade 5	Grade 6	Total
Designing Parachutes	Control	112	132	304	-	548
	Test	346	379	744	121	1590
Designing Solar Ovens	Control	159	7	206	134	506
	Test	162	167	358	80	767

The *Designing Parachutes* test sample included a smaller proportion of Hispanic students and a larger proportion of white students, but otherwise was similar to the control sample (Table 3). The control samples included a larger proportion of minority students than the test sample. The test sample for *Designing Solar Ovens* included slightly fewer black students and Hispanic students, and more white and Asian students, than the control sample.

Table 3. Sample Size by Assessment, by Race/Ethnicity

		Black	Asian	Hispanic	White	Other	Total
Designing Parachutes	Control	6.5%	5.9%	25.3%	58.1%	4.2%	525
	Test	6.9%	4.4%	9.7%	73.8%	5.1%	1329
Designing Solar Ovens	Control	8.2%	6.8%	12.6%	66.9%	5.5%	438
	Test	6.3%	9.7%	9.0%	68.3%	6.9%	715

The control samples for both assessments included a slightly higher proportion of students receiving free or reduced-price lunch. For *Designing Solar Ovens*, this difference was not significant (Pearson Chi-Square p=.091); for *Designing Parachutes*, it was significant (Pearson Chi-Square p=.012).

Table 4. Sample Size by Assessment, by % Students Receiving Free or Reduced Lunch

		%FRL (of N reporting)	Pearson χ^2 p=	N Reporting	N Missing	Total N
Designing Parachutes	Control	28.2%	**.012**	479	46	525
	Test	22.4%		1326	3	1329
Designing Solar Ovens	Control	26.3%	.091	418	20	438
	Test	21.5%		479	236	715

"Designing Solar Ovens" Unit: Summative Evaluation

The EiE unit *Green Engineering: Designing Solar Ovens* focuses on green engineering, the science of energy and energy transfer, and how green engineers use their knowledge about energy, resources, and local consumers to create environmentally friendly technologies. Lesson 1 of the unit explores the story of Lerato, a young girl in Botswana, who learns about solar ovens from an older girl from her village who is studying to be a green engineer. Lerato uses the solar oven to cook food for her family, so that she and her brothers and sister do not have to collect firewood each night for the next day's fire. This lesson introduces students to the basic concepts and vocabulary of green engineering, as well as the basic concepts of heat energy, thermal conductors and insulators, and heat transfer. Lesson 2 teaches students about the "life cycle assessment" used by engineers to evaluate the environmental impact of various technologies. Children follow the life cycle of a glass bottle and think about the differences for the environment when the bottle is reused, recycled, or simply thrown out. In Lesson 3, students learn about thermal conductors versus thermal insulators and conduct experiments to test the thermal properties of different materials. This leads them into Lesson 4, in which they use the results of those experiments to design and test their own solar ovens.

"Designing Solar Ovens": Assessment Instrument

The *Designing Solar Ovens* assessment was created and piloted in Year 5 (2007-2008) of the EiE project with 16 questions. Based on responses from pilot classroom students, a number of questions were cut or revised. The revised assessment, 9 questions long, was distributed to field test teachers during the 2008-2009 school year (Year 6 of the EiE project). Only one version of the *Designing Solar Ovens* assessment was designed for grades 3-5; as the content of the unit was deemed too advanced for grade 2 students, a grade 2 version of the *Designing Solar Ovens* assessment was not developed.

Table 5. Designing Solar Ovens Questions

Question #	Question Type	Question and Answer Text
1R	Green Engineering	Sophie practices green engineering. She is helping to design water bottles in a factory. Which would she be MOST concerned with? [how to make the water bottles a more popular color / **how to make the water bottles easy to reuse or recycle** / how to design machines to make lots of water bottles / how to make the water bottles easy to carry]
6R	Green Engineering	Too much trash is a problem around the world. How could the trash problem be improved in a way that is MOST friendly to the environment? [Recycle more trash. / Send trash into space. / **Find ways to make less trash.** / Bury trash under lots of dirt to make parks.]
16R	Green Engineering	Which of the following are resources that are needed to make cookies? [flour, sugar, and other cookie ingredients / people who grow the wheat in the cookies / energy used to bake the cookies / **all of the above**]
3	Science / Engineering (Insulation)	Ken has an insulated lunch bag. He is using it to keep his yogurt cool. How does it work? [It keeps cold inside / It makes things cool. / **It allows heat to pass through it very slowly.** / It does not allow any heat to pass through.]
4	Science / Engineering (Insulation)	The next day, Ken wants to bring hot soup to work. Can he use the same lunch bag to keep his soup warm? [Yes, but only if he turns the bag inside out. / **Yes, an insulated lunch bag can keep hot foods warm and cold foods cool.** / No, because it only works to keep cold foods cool. / Not any more, because he just used this bag to keep foods cool.]
10	Science (Heat Energy)	If you put a wooden spoon and a metal spoon into a pot of boiling water the metal spoon will feel much warmer. Why? [**Heat moves more easily through metal than through wood.** / Heat moves more easily through wood than through metal. / The metal spoon was warmer than the wooden spoon before it went into the pot. / Wood holds cold really well.]
13R	Science (Heat Energy)	Juan poured hot soup into a cool bowl. Soon his bowl was too hot to carry. Which of the following also happened? [The soup got cooler than the bowl. / The bowl transferred cold to the air. / **The soup got cooler than it had been.** / The bowl transferred cold to the soup.]
14R	Engineering a Solar Oven	It is a sunny day but Raj's solar oven is not getting hot enough. What could he do to make it work better? [Insulate the box better. / Turn the oven so it catches more sunlight. / Tape up all the holes in the box so no air can get in or out. / **All of these ideas would work**.]
15R	Engineering a Solar Oven	Which of the following BEST describes how a solar oven works? [**It reflects light energy from the Sun onto a cooking pot.** / It transforms energy from the Sun into electrical energy. / It changes light energy into microwaves to heat things up. It takes cold energy away from food and replaces it with heat energy.]

The text of the nine multiple-choice questions on the field test version of the *Designing Solar Ovens* assessment are shown in Table 5, together with text of answer choices. The correct answer for each question is bolded. Each question was categorized before analysis into one of four categories or learning objectives addressed by the unit: Green Engineering, Insulation, Science (Heat Energy), and Engineering a Solar Oven.

"Designing Solar Ovens" Assessment: Sample Size and Demographic Distribution

Assessments were administered to students in 5 states: California, Massachusetts, Colorado, Minnesota, and New Hampshire. Massachusetts students represented more than one third of the sample. Grade 5 students completed just under half of the total number of surveys administered. A total of 1273 students took both a pre- and a post-assessment; 767 of them were EiE (test) students and 506 were control (Table 6).

Table 6. *Designing Solar Ovens* Assessment: Sample Size by Grade, by State

		CA	CO	MA	MN	NH	Total
Grade 3	Control	109	44	6	0	-	159
	Test	84	46	19	13	-	162
Grade 4	Control	-	-	7	0	-	7
	Test	-	54	30	83	-	167
Grade 5	Control	-	-	122	47	37	206
	Test	-	79	172	70	37	358
Grade 6	Control	-	-	0	134	-	134
	Test	-	-	80	0	-	80
Total	Control	109	44	135	181	37	506
	Test	84	179	301	166	37	767

Extrapolating from the number of students whose race/ethnicity was reported (N reporting = 1153/1273 or 91%), the majority of the students in the sample were White/Caucasian (67.7%). Hispanic students were the next largest group (10.3%), followed by Asian students (8.6%). Black students represented 7.0% of the sample and those identifying as "other" were 6.3% of the sample (Table 7).

Table 7. Designing Solar Ovens Assessment: Sample Size by Race/Ethnicity

	Black	Asian	Hispanic	White	Other	Total
Control	36 / 8.2%	30 / 6.8%	55 / 12.6%	293 / 66.9%	24 / 5.5%	438
EiE	45 / 6.3%	69 / 9.7%	64 / 9.0%	488 / 68.3%	49 / 6.9%	715
Total	81	99	119	781	73	1153
Total (Percent)	7.0%	8.6%	10.3%	67.7%	6.3%	100%

Girls and boys each made up approximately half of the sample (Table 8). Free or Reduced Lunch status was reported for 95% of the control sample but only 67% of the test sample; of these, 103 test students (21.5% of students for whom data was reported) received "Free Lunch" or "Reduced Lunch" from the National School Lunch Program, while 110 control students (26.3% of those not missing data) received free or reduced lunch. This difference was significant (p<.000).

Table 8. Designing Solar Ovens Assessment: Sample Size by Free or Reduced Lunch

	FRL	% FRL (of N Reporting)	N Reporting	% of Sample Reporting	Total N
Control	110	26.3%	418	95.4%	438
EiE	103	21.5%	479	67.0%	715
Total	213	23.7%	897	77.8%	1153

“Designing Solar Ovens” Assessment: Analysis and Findings

The field test version of the *Designing Solar Ovens* assessment includes 9 multiple-choice questions. For each question, student responses were converted to a score of correct (1) / incorrect (0). These scores were summed, and the resulting overall score analyzed for normality, internal-consistency reliability and exploratory factor analysis (principal components with oblimin rotation) in SPSS. Reliability measured on the post-assessments only was poor (Cronbach’s alpha =.353) with little prospect for improvement by removing questions. It was hoped that principle components analysis would discriminate between science and engineering questions, but the questions were too few and too diverse in topic. The various indicators of factorability were fair to poor; though Bartlett’s Test of Sphericity was highly significant (p<.000) and the KMO Measure of Sampling Adequacy was acceptable (KMO=.614), observed correlations were quite low and many residuals quite large. Three components were extracted with eigenvalues greater than 1.0; however the scree plot indicated only one. The three components did not correspond to our theory of which questions would go together (categorized as science questions / green engineering questions / insulation questions / solar ovens questions—see above).

Given the poor prospects for combining items into a reliable scale, the decision was made to analyze student responses to items individually. Within-group differences from pre- to post-assessment were tested for significance using McNemar’s Test of Symmetry as the most appropriate for dichotomous data (correct/incorrect), especially pre/post data. Between-group differences on the pre-assessment and on the post-assessment were tested for significance using the phi coefficient, a chi-square based measure of association for use with dichotomous data.

Student responses are summarized in Table 9. For items listed in the left-most column of Table 9, overall performance (measured by percentage of students answering correctly) and grade-level performance are given for both the EiE (test) sample and the control sample. Exact significance for within-group differences is reported under “McNemar p=”. P-values significant at p<.05 or below are highlighted in bold. Between-group significance for control versus test on the pre-survey and on the post-survey is given in the final two columns; this was analyzed using the phi coefficient. Significant values are marked in bold.

Analysis of students’ completed assessments for *Designing Solar Ovens* resulted in several significant and notable findings (see Figure 2, Figure 3, and Table 9). Comparing students in the EiE (test) group with their control counterparts, we found that there were no significant differences between the two groups’ performances on the pre-assessments for all questions except two, on which control students performed significantly better than test students (questions 13R and 16R). However the EiE group performed significantly better than the control group on five of the nine post-assessment questions, including 2 of 3 Green Engineering questions, both Solar Ovens questions, and one Insulation question. EiE (test) students improved significantly on 8 of the 9 questions (all but question 13R, a Heat Energy question). Control students improved significantly on 6 of the 9 questions. Overall, EiE students appear to have improved much more than control students on 6 of the 9 questions on the post-assessment—including all engineering-content questions and one Insulation question.

Figure 2. ***Designing Solar Ovens*** **Assessment: Engineering Questions**

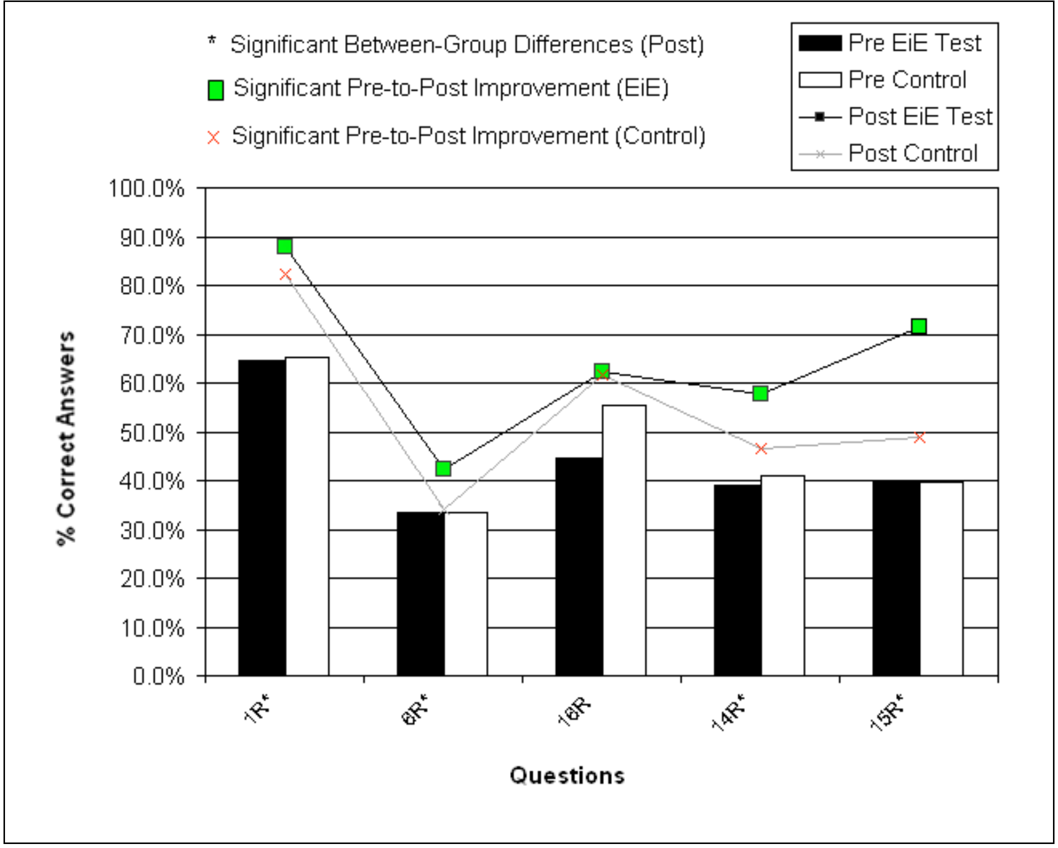


Figure 3. ***Designing Solar Ovens*** **Assessment: Science Questions**

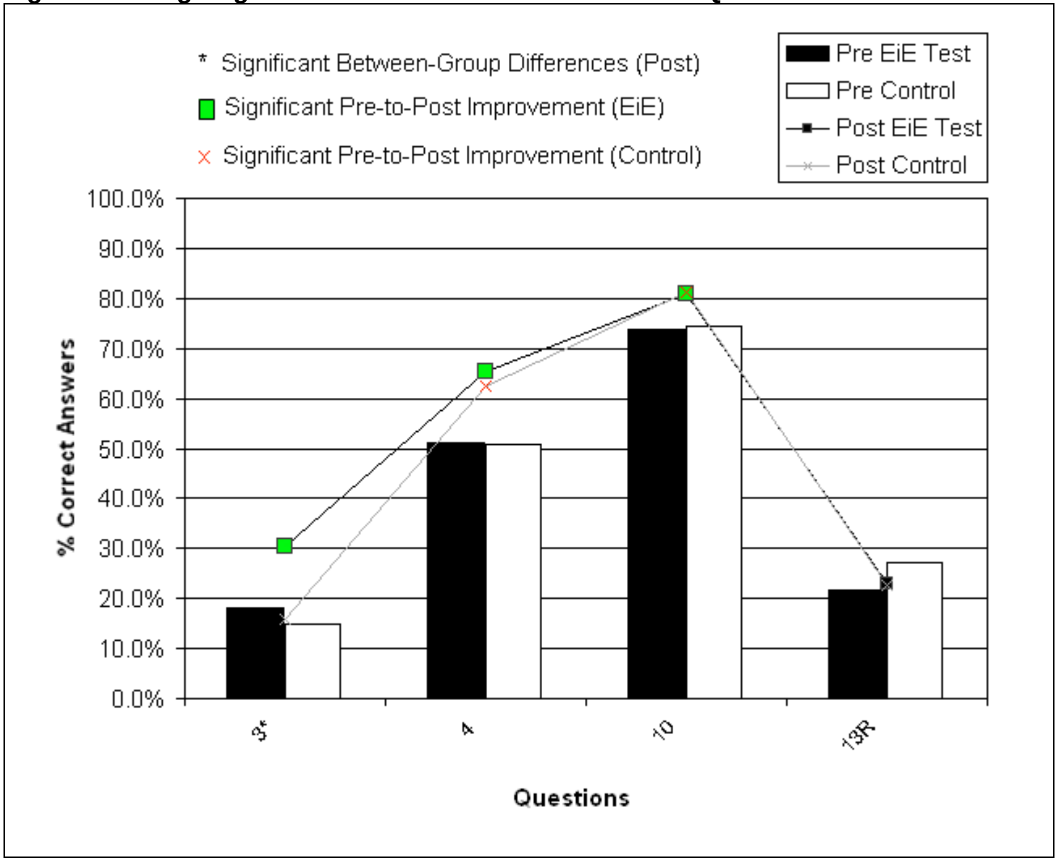


Table 9. Designing Solar Ovens Assessment: Results (EiE Vs. Control)

		Within-Group Differences (Pre vs. Post)								**Test/Control Differences**	
		EiE (Test)				**Control**				**PRE**	**POST**
Q#	**Group**	**N**	**Pre**	**Post**	**p=**	**N**	**Pre**	**Post**	**p=**	**p=**	**p=**
1R (GE)	Total	760	64.6%	88.7%	**.000**	459	65.1%	82.4%	**.000**	.849	**.002**
	Grade 3	159	47.2%	81.8%	**.000**	156	47.4%	67.3%	**.000**	.962	**.003**
	Grade 4	166	62.7%	93.4%	**.000**	Insufficient N to report				.637	.435
	Grade 5	355	70.1%	87.9%	**.000**	163	77.3%	87.7%	**.006**	.091	.959
	Grade 6	80	78.8%	96.3%	**.001**	133	70.7%	93.2%	**.000**	.195	.355
6R (GE)	Total	762	33.6%	42.3%	**.000**	464	33.6%	34.1%	.935	.993	**.004**
	Grade 3	161	42.9%	51.6%	.098	157	33.8%	38.2%	.419	.095	.017
	Grade 4	166	38.6%	47.6%	.077	Insufficient N to report				.819	.323
	Grade 5	355	28.7%	38.0%	**.003**	167	39.5%	29.9%	**.029**	**.014**	.072
	Grade 6	80	26.3%	31.3%	.454	133	25.6%	34.6%	.096	.912	.617
16R (GE)	Total	745	44.6%	62.3%	**.000**	460	55.4%	61.7%	**.008**	**.000**	.850
	Grade 3	158	13.9%	41.1%	**.000**	154	20.8%	34.4%	**.002**	.110	.221
	Grade 4	160	50.6%	68.1%	**.000**	Insufficient N to report				.253	.325
	Grade 5	349	51.6%	67.3%	**.000**	167	67.1%	62.9%	.349	**.001**	.317
	Grade 6	78	62.8%	70.5%	.263	132	82.6%	90.9%	**.043**	**.001**	**.000**
3 (Ins.)	Total	764	18.1%	30.5%	**.000**	468	15.0%	16.0%	.691	.158	**.000**
	Grade 3	162	20.4%	34.6%	**.004**	159	20.8%	25.2%	.443	.932	.066
	Grade 4	166	25.3%	30.1%	.341	Insufficient N to report				.846	.368
	Grade 5	356	14.9%	28.4%	**.000**	168	14.3%	14.9%	1.000	.856	**.001**
	Grade 6	80	12.5%	32.5%	**.002**	134	8.2%	6.7%	.754	.307	**.000**
4 (Ins.)	Total	764	51.0%	65.4%	**.000**	467	50.7%	62.3%	**.000**	.919	.266
	Grade 3	161	31.7%	53.4%	**.000**	158	30.4%	50.6%	**.000**	.802	.619
	Grade 4	165	49.7%	63.0%	**.009**	Insufficient N to report				.700	.221
	Grade 5	358	55.3%	68.2%	**.000**	168	49.4%	56.5%	.134	.206	**.010**
	Grade 6	80	73.8%	82.5%	.167	134	76.1%	82.1%	.200	.698	.939
10 (HE)	Total	757	73.7%	81.0%	**.000**	455	74.3%	81.1%	**.004**	.826	.958
	Grade 3	157	58.0%	66.9%	.114	154	61.0%	65.6%	.392	.580	.809
	Grade 4	165	68.5%	77.6%	.058	Insufficient N to report				.156	.611
	Grade 5	355	81.4%	85.1%	.154	162	79.0%	87.7%	**.024**	.522	.434
	Grade 6	80	81.3%	97.5%	**.001**	132	85.6%	90.9%	.189	.402	.061
13R (HE)	Total	761	21.6%	23.0%	.534	463	27.2%	22.7%	.085	**.024**	.898
	Grade 3	162	27.8%	21.0%	.193	157	29.9%	20.4%	.058	.671	.894
	Grade 4	165	24.2%	20.6%	.519	Insufficient N to report				.265	.161
	Grade 5	354	17.8%	24.0%	**.043**	167	24.6%	23.4%	.880	.072	.869
	Grade 6	80	20.0%	27.5%	.377	132	26.5%	23.5%	.608	.282	.513
14R (SO)	Total	766	38.9%	57.6%	**.000**	469	40.9%	46.5%	**.041**	.478	**.000**
	Grade 3	161	24.8%	32.3%	**.088**	159	24.5%	32.7%	.105	.948	.938
	Grade 4	167	38.3%	65.9%	**.000**	Insufficient N to report				.809	.211
	Grade 5	358	43.6%	64.5%	**.000**	169	37.3%	40.8%	.512	.171	**.000**
	Grade 6	80	47.5%	60.0%	.132	134	64.9%	70.1%	.311	**.012**	.128
15R (SO)	Total	765	39.9%	71.4%	**.000**	467	39.6%	49.0%	**.002**	.929	**.000**
	Grade 3	162	34.6%	70.4%	**.000**	159	32.1%	44.0%	**.027**	.636	**.000**
	Grade 4	167	36.5%	67.1%	**.000**	Insufficient N to report				.229	.810
	Grade 5	356	44.9%	74.7%	**.000**	167	48.5%	46.1%	.694	.446	**.000**
	Grade 6	80	35.0%	67.5%	**.000**	134	38.8%	57.5%	**.001**	.578	.145

Figure 4. ***Designing Solar Ovens*** **Assessment: Engineering Questions by Gender**

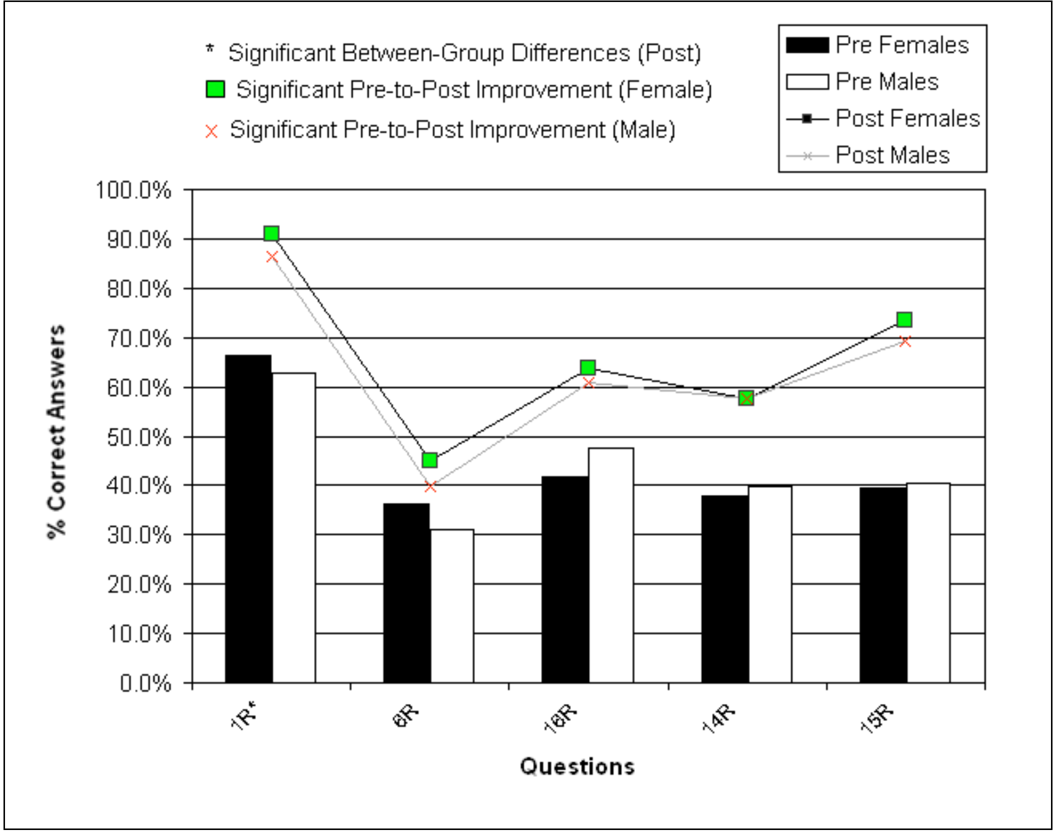


Figure 5. ***Designing Solar Ovens*** **Assessment: Science Questions by Gender**

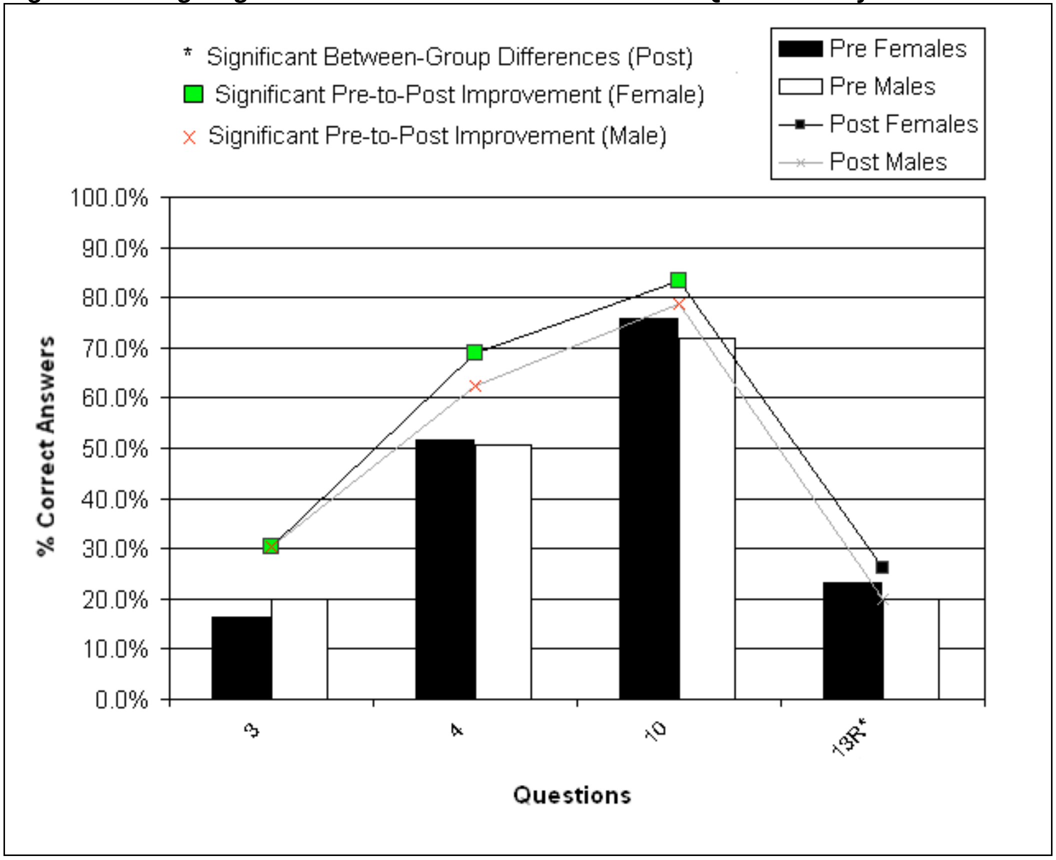


Table 10. Designing Solar Ovens Assessment: Gender Differences (EiE only)

	Within-Group Differences (Pre vs. Post)								Male / Female Differences	
	Female				Male				PRE	POST
Q#	N	Pre	Post	p=	N	Pre	Post	p=	p=	p=
1	372	66.4%	91.1%	**.000**	388	62.9%	86.3%	**.000**	.312	**.037**
6R	373	36.2%	45.0%	**.005**	388	31.2%	39.7%	**.004**	.144	.135
16R	367	41.7%	63.8%	**.000**	377	47.5%	61.0%	**.000**	.112	.439
3	375	16.3%	30.4%	**.000**	388	19.8%	30.4%	**.000**	.199	.997
4	375	51.5%	68.8%	**.000**	388	50.8%	62.4%	**.000**	.848	.062
10	370	75.7%	83.2%	**.007**	387	71.8%	78.8%	**.012**	.230	.120
13R	373	23.3%	26.3%	.396	387	19.9%	19.9%	1.000	.251	**.037**
14R	375	37.9%	57.6%	**.000**	390	39.7%	57.7%	**.000**	.594	.979
15R	376	39.4%	73.4%	**.000**	388	40.5%	69.3%	**.000**	.756	.213

Comparing male and female EiE students (see Figure 4, Figure 5, and Table 10), both genders showed significant pre-to-post assessment improvements on all questions except 13R (about Heat Energy). Girls performed significantly better than boys on two questions on the post-assessment: question 1(about Green Energy) and question 13R.

Table 11. Designing Solar Ovens Assessment: Free/Reduced Lunch Differences (EiE only)

	Within-Group Differences (Pre vs. Post)								FRL / Not FRL Differences	
	Free or Reduced Lunch				Not Free or Reduced Lunch				PRE	POST
Q#	N	Pre	Post	p=	N	Pre	Post	p=	p=	p=
1	103	48.5%	83.5%	**.000**	372	70.7%	95.7%	**.000**	**.000**	**.000**
6R	102	27.5%	38.2%	.099	375	38.9%	47.2%	**.008**	**.033**	.107
16R	100	33.0%	59.0%	**.000**	365	44.7%	60.3%	**.000**	**.036**	.818
3	103	20.4%	31.1%	.099	373	16.9%	33.0%	**.000**	.410	.715
4	103	46.6%	53.4%	.281	374	55.3%	74.6%	**.000**	.115	**.000**
10	102	71.6%	80.4%	.136	370	72.2%	83.2%	**.000**	.906	.501
13R	103	21.4%	21.4%	1.000	374	22.2%	23.0%	.860	.857	.725
14R	103	31.1%	50.5%	**.002**	375	39.2%	59.2%	**.000**	.131	.113
15R	103	41.7%	74.8%	**.000**	375	39.5%	70.7%	**.000**	.676	.415

Students who did not receive free or reduced lunch (the N-FRL students) significantly out-performed students receiving free or reduced lunch (the FRL students) on three pre-assessment questions and two post-assessment questions (see Figure 6, Figure 7, and Table 11). N-FRL students improved significantly on all questions except one; FRL students improved significantly on only four questions, though they did improve on the other questions on which N-FRL students improved. With only a fifth the sample receiving free or reduced lunch, significance is more difficult to measure for the FRL students.

Figure 6. ***Designing Solar Ovens*** **Assessment: Engineering Questions by FRL Status**

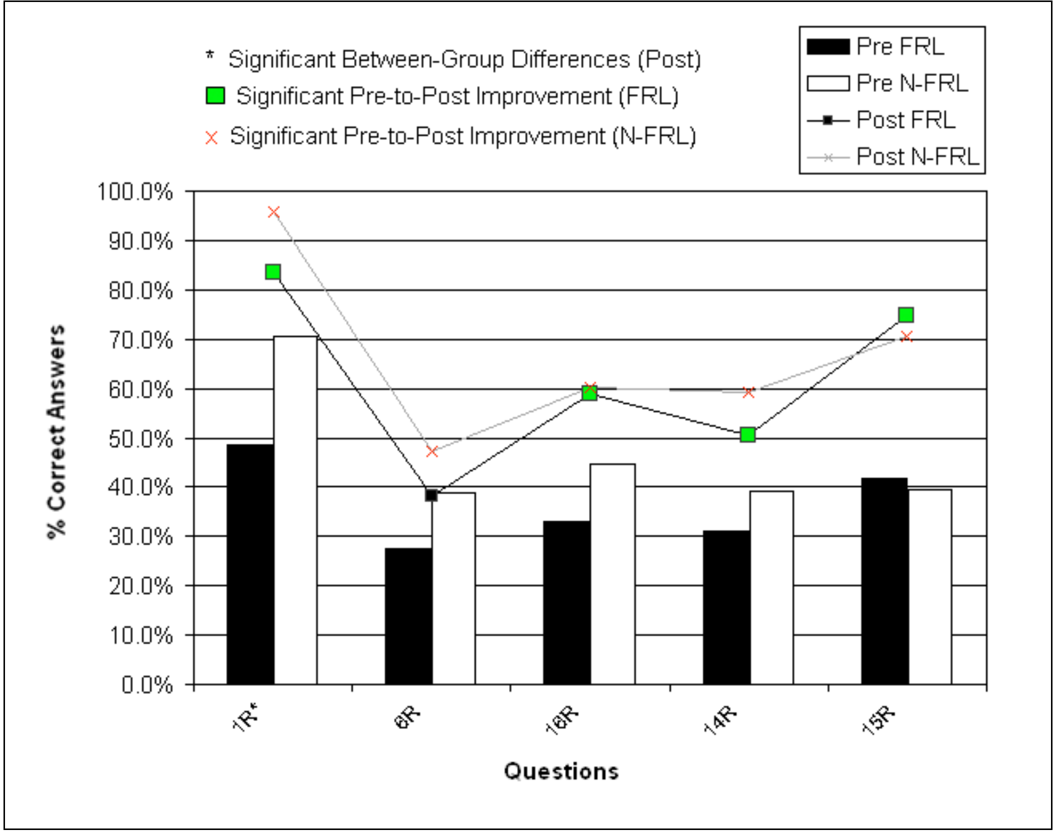
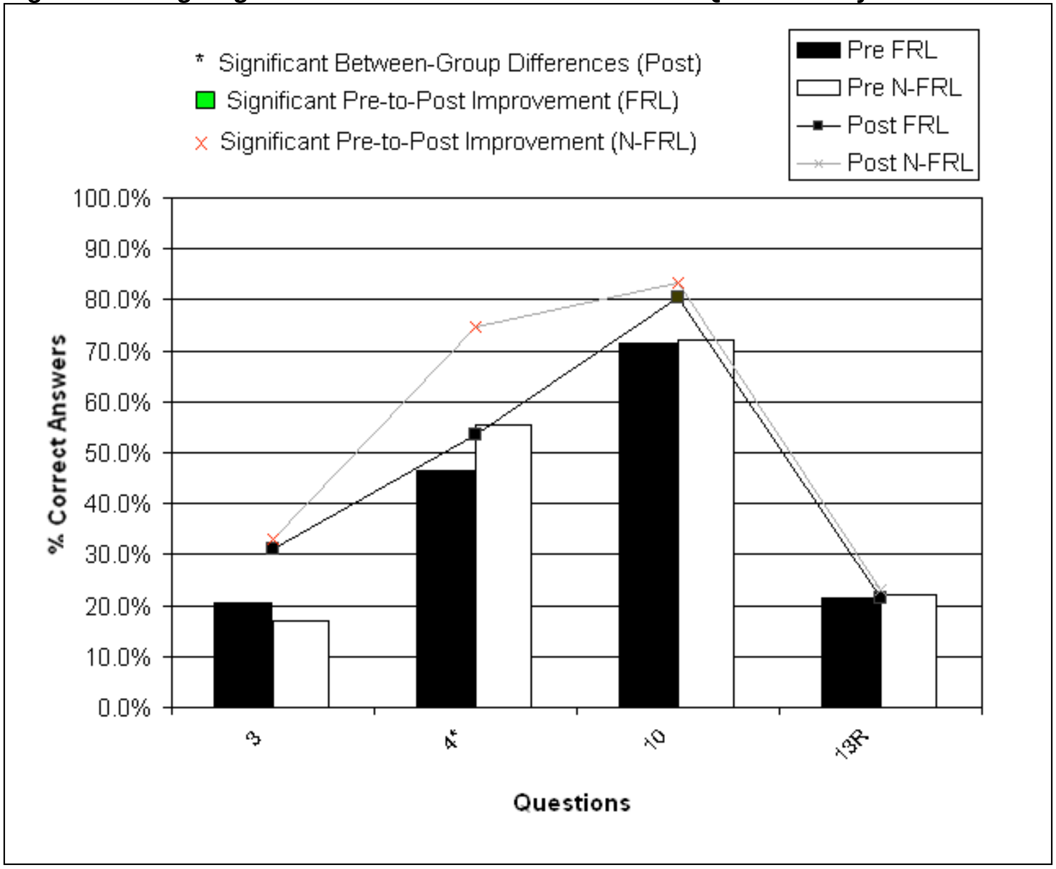


Figure 7. ***Designing Solar Ovens*** **Assessment: Science Questions by FRL Status**



“Designing Parachutes” Unit: Summative Evaluation

In the EiE unit *Designing Parachutes,* students are introduced to the field of aerospace engineering and how atmospheric properties and the concept of drag affect the design and operation of parachutes beginning with the discussion of a fictional story in Lesson 1. Students discuss and brainstorm the work of aerospace engineers and the properties of the planets and other celestial bodies as they determine how a fictional spacecraft will navigate the solar system (Lesson 2). They then explore the concept of drag by observing and analyzing a model that demonstrates how the thickness of an atmosphere affects how fast objects fall. Students also collect and analyze their own data relating to three variables of a parachute: suspension line length, canopy size, and canopy material (Lesson 3). Finally in Lesson 4, students utilize their data to inform the design of their own parachutes. Using the Engineering Design Process they plan, create, test, and improve their parachute designs to meet packing and speed criterion for a mission travelling to a planet with an atmosphere thinner than Earth’s.

“Designing Parachutes”: Assessment Instrument

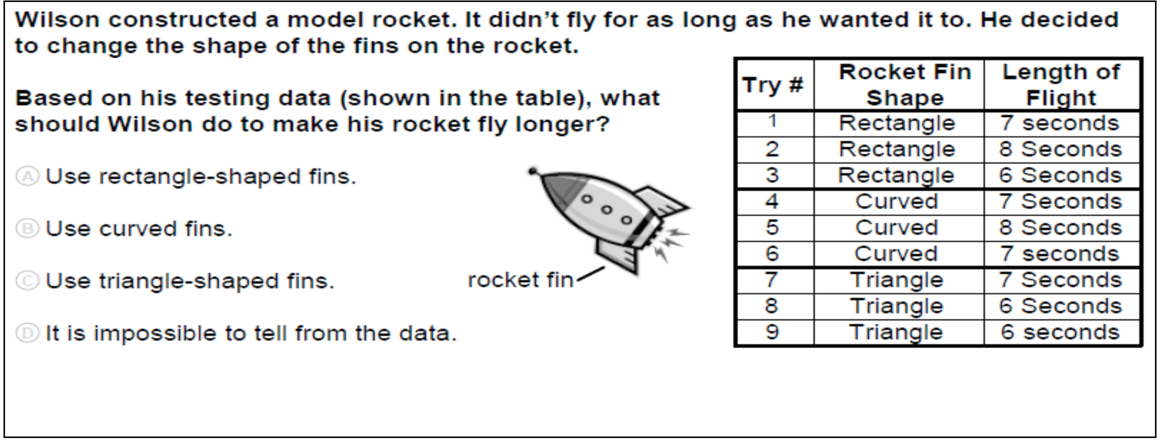
The Designing Parachutes assessment was designed and first tested during Year 5 of the EiE project (the 2007-2008 school year). Based on statistical analysis, it was revised during the summer of 2008, and a number of questions were dropped or changed. This “field test” version of the assessment, with ten questions, was administered to students during the 2008-2009 school year – Year 6 of the EiE project.

Table 12. Designing Parachutes Assessment: Questions (Text)

Question #	Question Type	Question Text
1	Engineering	The atmosphere on Mars is thinner than the atmosphere on Earth. How can Mr. Ino change a parachute that works well on Earth to help a robot land safely on Mars? [Use a parachute material that lets more air through. / Use a parachute that weighs more. / **Make the parachute larger.** / All of the above.]
7	Engineering	See Figure 1. [use rectangle-shaped fins / **use curved fins** / use triangle-shaped fins / it is impossible to tell from the data]
11R	Engineering	At work, an aerospace engineer might: [fly space shuttles. / study a rock from Mars. / fix airplane engines. / **figure out ways to help airplanes land safely.**]
13R	Engineering	See Figure 9. [**Make the canopy smaller.** / Make the suspension lines longer. / Drop the parachute from a lower spot. / Make the canopy out of a material that lets less air through.]
3	Science	On a planet the same size as Earth, where the atmosphere is thinner than it is on Earth, objects will: [**fall faster.** / fall slower. / fall at the same rate. / float.]
4	Science	David and Dipa took identical sheets of paper. They crumpled one into a ball and left one flat. They dropped them both at the same time. The crumpled paper ball dropped quickly straight down. The flat sheet of paper fell slowly and drifted to the ground. Why? [The crumpled paper ball weighed less than the flat sheet of paper. / The crumpled paper ball was more dense than the flat sheet of paper. / **More air got in the way of the flat sheet of paper and slowed it down.** / The air was denser near the flat sheet of paper and slowed it down.]
Questions 14d, 15c, 15g, and 9R were preceded by the directions: Which of the following statements are true? Mark “T” for true and “F” for false for each statement.		
9R	Engineering	A parachute will work on a planet with no atmosphere. [False]
14d	Science	Mars takes 365 days to circle the Sun. [False]
15c	Science	Only Earth has an atmosphere. [False]
15g	Science	Air can affect how things fall. [True]

The pre- and post-assessments were identical. Students were asked science questions about atmospheric properties and the effects of drag on falling objects; they were asked engineering questions about the design and operation of parachutes and the work of aerospace engineers. The questions are described below in Table 12, together with text of answer choices. The correct answer for each question is bolded. Each question was categorized before analysis into one of two categories: Science or Engineering.

Figure 8. Question 7, *Designing Parachutes*.



Wilson constructed a model rocket. It didn't fly for as long as he wanted it to. He decided to change the shape of the fins on the rocket.

Based on his testing data (shown in the table), what should Wilson do to make his rocket fly longer?

Ⓐ Use rectangle-shaped fins.

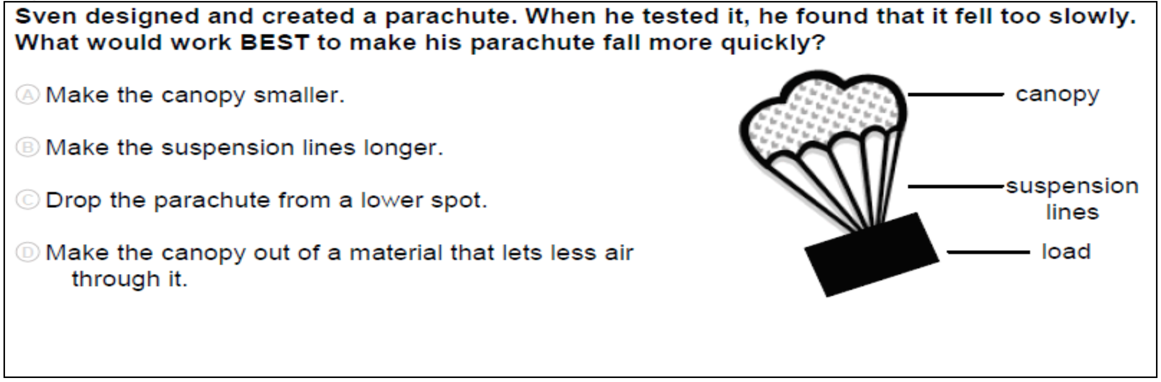
Ⓑ Use curved fins.

Ⓒ Use triangle-shaped fins.

Ⓓ It is impossible to tell from the data.

Try #	**Rocket Fin Shape**	**Length of Flight**
1	Rectangle	7 seconds
2	Rectangle	8 Seconds
3	Rectangle	6 Seconds
4	Curved	7 Seconds
5	Curved	8 Seconds
6	Curved	7 seconds
7	Triangle	7 Seconds
8	Triangle	6 Seconds
9	Triangle	6 seconds

Figure 9. Question 13R, *Designing Parachutes*.



Sven designed and created a parachute. When he tested it, he found that it fell too slowly. What would work BEST to make his parachute fall more quickly?

Ⓐ Make the canopy smaller.

Ⓑ Make the suspension lines longer.

Ⓒ Drop the parachute from a lower spot.

Ⓓ Make the canopy out of a material that lets less air through it.

The *Designing Parachutes* Assessment includes 6 multiple choice questions and 4 true/false questions. A future version of the assessment will have the true/false questions replaced by multiple-choice so that all questions have equal probability of correctness given guessing. However, it was decided for the purpose of this report, because of the limited number of questions used, to analyze all questions equivalently regardless of type.

"Designing Parachutes" Assessment: Sample Size and Demographic Distribution

The test sample included 1590 students who completed both a science unit on astronomy or the solar system as well as the EiE unit *Designing Parachutes* between the pre-assessment and post-assessment administrations (see Table 13). In addition, there were 548 control students. Control students completed only a science unit on astronomy or the solar system after the pre-assessment and before the post-assessment was administered. Both the control and test teachers were free to use whatever science curriculum they wished to complete the required science instruction during the field test period.

The sample was drawn from California, Colorado, Florida, Massachusetts, and Minnesota. Students in grades 3-6 were assessed. Girls made up 49.1% of the control sample, and 48.6% of the test sample, a difference which was not significant (Pearson Chi-Square p=.882).

Table 13. Designing Parachutes Assessment: Sample Size by Grade and by State

		CA	CO	FL	MA	MN	Total
Grade 3	Control	-	70	30	12	-	112
	Test	-	20	23	191	112	346
Grade 4	Control	-	77	15	40	-	132
	Test	-	113	-	164	102	379
Grade 5	Control	52	20	209	23	-	304
	Test	84	43	441	133	43	744
Grade 6	Control	-	-	-	-	-	-
	Test	-	-	-	95	26	121
Total	Control	52	167	254	75	-	548
	Test	84	176	464	583	283	1590

Table 14 describes the racial/ethnic distribution of the student population. White students made up more than half of the control sample and almost three-quarters of the test sample. Hispanic students made up the second largest population within each sample – 25.3% of the control sample, and 9.7% of the test sample. The test sample was 6.9% Black and 4.4% Asian, while the control sample was 6.5% Black and 5.9% Asian.

Table 14. Designing Parachutes Assessment: Sample Size by Race/Ethnicity

	Black	Asian	Hispanic	White	Other	Total
Control	34	31	133	305	22	525
EiE	92	59	129	981	68	1329
Total	126	90	262	1286	90	1854
Total (Percent)	6.8%	4.9%	14.1%	69.4%	4.9%	100.0%

Of the 1236 students in the test sample for whom information was available (see Table 15), 277 (22.4%) received free or reduced-price lunch (FRL) from the National School Lunch Program; 28.2% of the control sample (reporting) received free or reduced-price lunch. This difference was significant (Pearson Chi-Square p=.014).

Table 15. Designing Parachutes Assessment: Sample Size by Free or Reduced Lunch

	FRL	% FRL (of N Reporting)	N Reporting	% of Sample Reporting	Total N
Control	135	28.2%	479	91.2%	525
EiE	277	22.4%	1236	93.0%	1329
Total	412	24.0%	1715	92.5%	1854

"Designing Parachutes" Assessment: Analysis and Findings

The field test version of the *Designing Parachutes* assessment includes 6 multiple-choice questions and 4 True/False questions. For each question, student responses were converted to a score of correct (1) / incorrect (0). These scores were summed, and the resulting overall score analyzed for normality, internal-consistency reliability and exploratory factor analysis (principal components with oblimin rotation) in SPSS. Reliability measured on the post-assessments only was fair (Cronbach's alpha =.587). It was hoped that principle components analysis would discriminate between science and engineering questions, but only one component accounting for a significant percentage of the variance emerged. The various indicators of factorability were fair to poor; though Bartlett's Test of Sphericity was highly significant ($p<.000$) and the KMO Measure of Sampling Adequacy was acceptable (KMO=.741), observed correlations were quite low and some residuals quite large (greater than 0.2). Three components were extracted with eigenvalues greater than 1.0; however the scree plot indicated only one. Also, correlation and component matrices showed no discernable subscales could be made by dropping questions. The three components did not correspond to our theory of which questions would go together (categorized as science questions / engineering questions—see above). It was decided to use all questions together as a scale, despite the marginal reliability, and use results from an ANCOVA analysis of the all-item scale in comparison to chi-square analysis of the individual questions.

Before conducting an ANCOVA test on the all-item scale, regression analyses were conducted on a variety of independent variables by the dependent variable (post-assessment mean score) and covariate (pre-assessment mean score). No cases were found with significant Mahalanobis distance. Additionally the all-item scale was used to check the data for violations of homogeneity of regression and homogeneity of error variance in the independent variables. Grade 6 students were dropped from analysis because of violations of homogeneity of regression and homogeneity of error variance, as well as because of the small sample size and lack of control students (see Table 13). The Asian ethnic group was also dropped from analysis because of the same violations and small sample size.

Normality of the all-item scale was checked by visual inspection of the histograms for the scale. This inspection was deemed sufficient because of the large sample size. The dependent variable (the post-scale) was slightly negatively skewed, but this appeared to be due to a ceiling effect – the mean was close to the maximum score.

We conducted an ANCOVA test with the post all-item scale as dependent variable, and the pre all-item scale as covariate. Independent variables in the model were sample group (test or control), gender, and ethnic group (Levene's Test of Equality of Error Variances was not significant; $p=.051$). There were no significant interactions between any of these variables. The variable for free/reduced-price lunch was not included in the final ANCOVA model because it made error variance across students unequal. After adjusting for pretest scores, there were significant effects of test/control ($F(1,1519) = 55.677$, $p<.000$, partial $\eta^{\wedge 2} =.035$). Adjusted mean post all-item scale scores suggest that students participating in the EiE curriculum in conjunction with science instruction achieved increased scores (EiE adjusted mean = 6.830) on the engineering & science assessment compared with the control group (control adjusted mean = 5.634).

Analysis of individual questions of the *Designing Parachutes* assessment corroborates these findings. Overall, students in the EiE group made significant gains in pre-to-post assessment performance that were larger in magnitude than those obtained by students in the control group for most questions. For 7 out of the 10 assessment questions there were no significant differences in the pre-assessment scores between the EiE and control groups, but the EiE group had post-assessment scores that were significantly higher than those for the control group for 6 of these questions. Therefore, having started at indistinguishable pre-assessment scores, EiE students improved significantly more and performed significantly better than their control counterparts on three science questions (q3, q15c, and q15g) and three engineering questions (q11R, q7, and q13R) (see Figure 10, Figure 11, and Table 16).

Additionally, analyses of one additional engineering question (q1) and one additional science question (q4) showed that EiE students performed significantly better on the post-assessment than control students. However, both EiE and control showed significant within-group improvements from pre-to post-assessment, and the EiE group also performed significantly better on the pre-assessment than control on these two questions.

Figure 10. ***Designing Parachutes*** **Assessment: Engineering Questions**

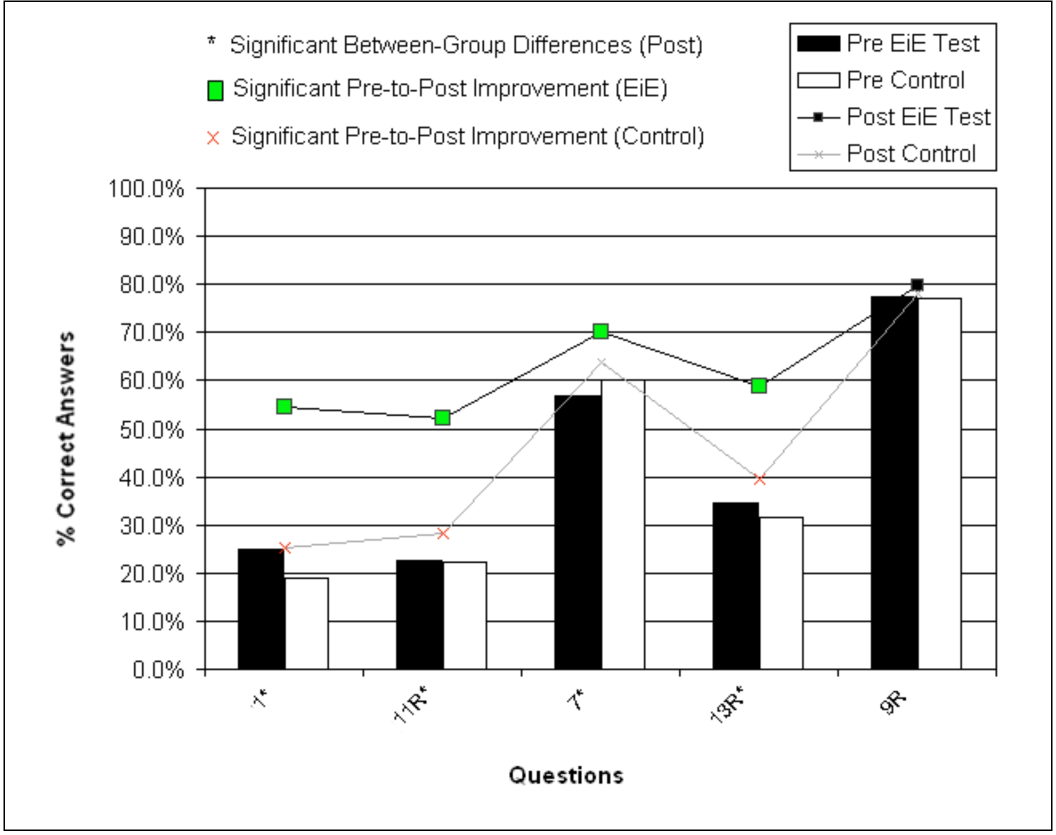


Figure 11. ***Designing Parachutes*** **Assessment: Science Questions**

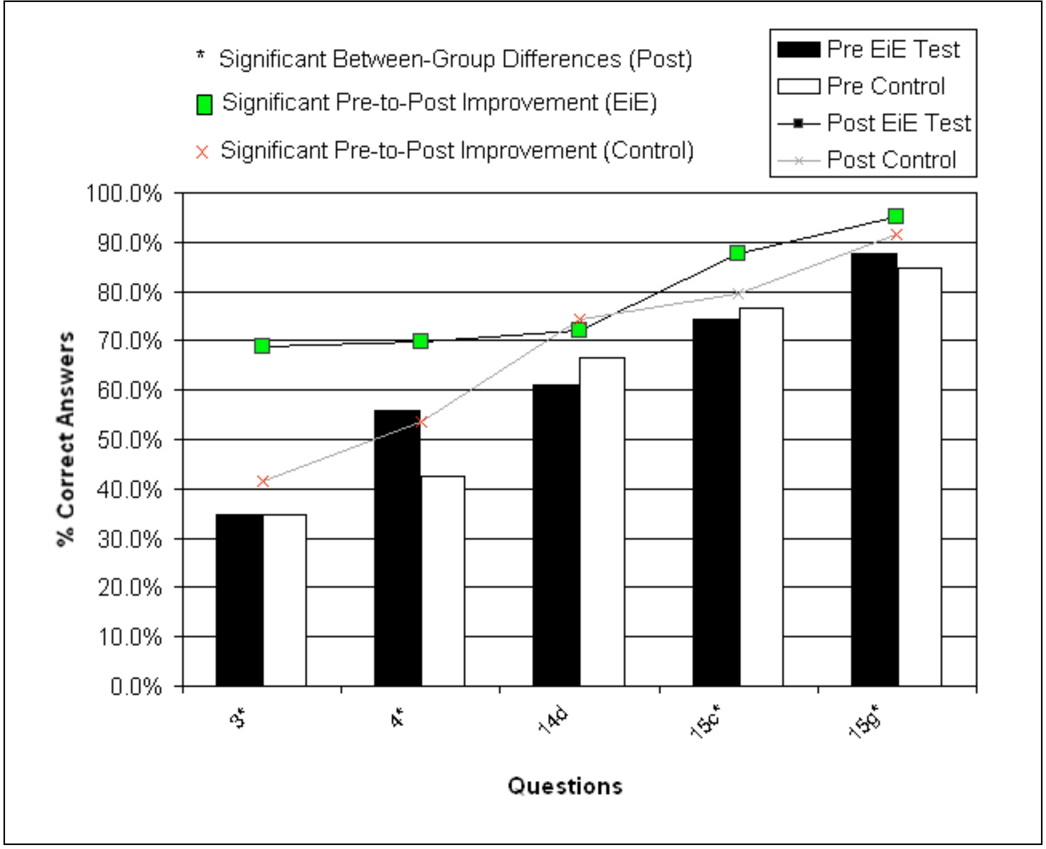


Table 16. Designing Parachutes Assessment: Results (EiE Vs. Control)*

		Within-Group Differences (Pre vs. Post)								**Test / Control Differences**	
		EiE (Test)				**Control**				**PRE**	**POST**
Q#	**Group**	**N**	**Pre**	**Post**	**p=**	**N**	**Pre**	**Post**	**p=**	**p=**	**p=**
1 (Eng)	Total	1174	25.0%	54.4%	**.000**	483	18.8%	25.3%	**.010**	**.007**	**.000**
	Grade 3	244	24.2%	56.1%	**.000**	100	17.0%	19.0%	.832	.145	**.000**
	Grade 4	348	27.0%	58.3%	**.000**	124	25.0%	37.1%	**.040**	.663	**.000**
	Grade 5	582	24.1%	51.4%	**.000**	259	16.6%	22.0%	.114	**.016**	**.000**
7 (Eng)	Total	1177	56.7%	70.2%	**.000**	490	60.2%	63.9%	.210	.183	**.012**
	Grade 3	246	50.8%	65.4%	**.001**	105	50.5%	56.2%	.504	.954	.101
	Grade 4	351	54.1%	70.7%	**.000**	126	57.1%	65.9%	.144	.560	.318
	Grade 5	580	60.7%	71.9%	**.000**	259	65.6%	66.0%	1.000	.172	.086
11R (Eng)	Total	1166	22.5%	52.2%	**.000**	478	22.2%	28.2%	**.021**	.897	**.000**
	Grade 3	240	22.5%	52.9%	**.000**	100	19.0%	25.0%	.345	.474	**.000**
	Grade 4	349	22.1%	57.9%	**.000**	123	22.0%	28.5%	.291	.979	**.000**
	Grade 5	577	22.7%	48.5%	**.000**	255	23.5%	29.4%	.105	.794	**.000**
13R (Eng)	Total	1168	34.4%	58.9%	**.000**	484	31.4%	39.7%	**.002**	.238	**.000**
	Grade 3	243	18.5%	56.8%	**.000**	105	20.0%	26.7%	.248	.746	**.000**
	Grade 4	346	33.8%	61.0%	**.000**	125	34.4%	42.4%	.193	.906	**.000**
	Grade 5	579	41.5%	58.5%	**.000**	254	34.6%	43.7%	**.015**	.064	**.000**
3 (Sci)	Total	1169	34.7%	68.9%	**.000**	484	34.9%	41.7%	**.017**	.942	**.000**
	Grade 3	244	25.8%	69.7%	**.000**	102	32.4%	43.1%	.126	.216	**.000**
	Grade 4	347	33.7%	71.2%	**.000**	123	36.6%	43.9%	.222	.565	**.000**
	Grade 5	578	39.1%	67.3%	**.000**	259	35.1%	40.2%	.213	.274	**.000**
4 (Sci)	Total	1177	55.7%	69.9%	**.000**	491	42.6%	53.6%	**.000**	**.000**	**.000**
	Grade 3	244	51.6%	66.8%	**.000**	104	48.1%	53.8%	.418	.543	**.022**
	Grade 4	353	62.9%	73.1%	**.001**	127	40.2%	59.8%	**.000**	**.000**	**.005**
	Grade 5	580	53.1%	69.3%	**.000**	260	41.5%	50.4%	**.017**	**.002**	**.000**
9R (Eng)	Total	1170	77.3%	79.8%	.135	481	77.1%	78.0%	.796	.953	.395
	Grade 3	242	80.6%	76.4%	.308	103	75.7%	81.6%	.345	.310	.295
	Grade 4	351	77.2%	80.1%	.387	123	74.8%	75.6%	1.000	.587	.298
	Grade 5	577	75.9%	81.1%	**.035**	255	78.8%	77.6%	.813	.359	.249
14d (Sci)	Total	1176	61.1%	72.2%	**.000**	485	66.6%	74.2%	**.003**	**.036**	.397
	Grade 3	246	52.0%	50.4%	.755	105	61.0%	84.8%	**.000**	.124	**.000**
	Grade 4	352	54.3%	72.4%	**.000**	123	56.1%	65.0%	.161	.725	.121
	Grade 5	578	69.2%	81.3%	**.000**	257	73.9%	74.3%	1.000	.166	**.022**
15c (Sci)	Total	1172	74.4%	87.7%	**.000**	480	76.5%	79.4%	.265	.381	**.000**
	Grade 3	241	66.4%	84.6%	**.000**	104	76.0%	70.2%	.377	.077	**.002**
	Grade 4	351	75.8%	86.6%	**.000**	122	79.5%	78.7%	1.000	.402	**.037**
	Grade 5	580	76.9%	89.7%	**.000**	254	75.2%	83.5%	**.015**	.595	**.012**
15g (Sci)	Total	1176	87.6%	95.2%	**.000**	483	84.7%	91.5%	**.000**	.113	**.004**
	Grade 3	245	83.3%	94.3%	**.000**	104	83.7%	91.3%	.077	.929	.311
	Grade 4	352	89.2%	95.5%	**.002**	123	86.2%	93.5%	.078	.367	.393
	Grade 5	579	88.4%	95.3%	**.000**	256	84.4%	90.6%	**.029**	.106	**.009**

*excluding gr6 and Asian ethnic grp from sample (because of violations of homogeneity of regression and error variance on the all_scale)

Figure 12. ***Designing Parachutes*** **Assessment: Engineering Questions by Gender**

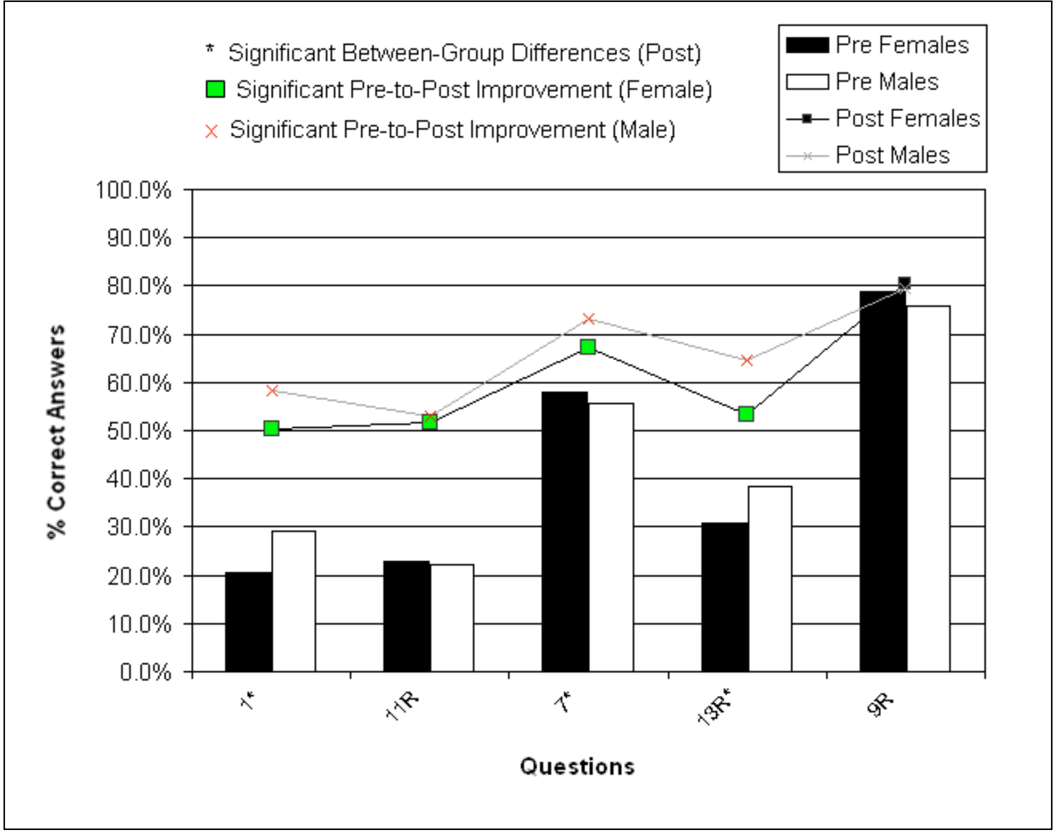
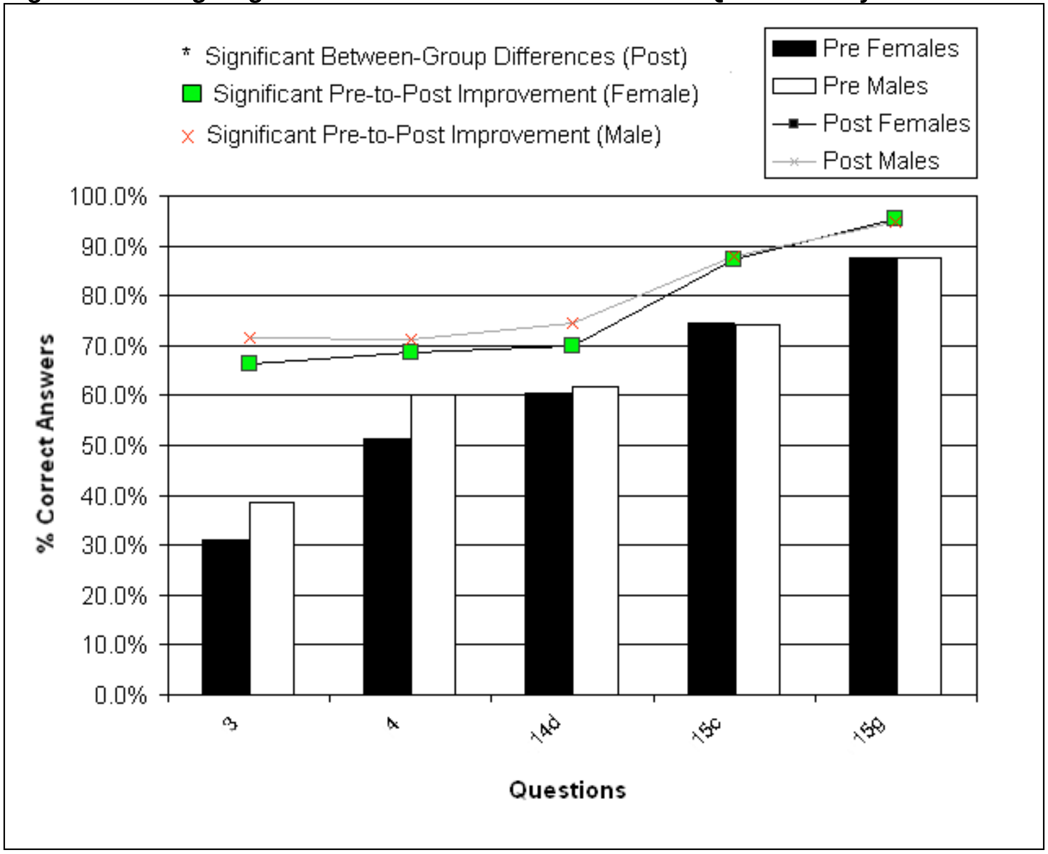


Figure 13. ***Designing Parachutes*** **Assessment: Science Questions by Gender**



For gender, there were also significant effects in the ANCOVA model after adjusting for pretest scores (F(1,1519) = 5.766, p=.016, partial $\eta^{\wedge 2}$ = .004). Adjusted mean post all-item scale scores suggest that males overall performed better on the post-assessment (male adjusted mean = 6.425) than females overall (female adjusted mean = 6.040). The interaction effect between gender and test/control comparison was not significant (test/control*gender F(1,1519) = .000, p=.993, partial $\eta^{\wedge 2}$ = .000).

However, nonparametric tests conducted on each of the *Designing Parachutes* assessment questions (Figure 12, Figure 13, and Table 17) found that, within the test group at least, males and females did not perform significantly differently on most questions. For all questions but q9R, both females and males showed significant pre-to-post performance improvements that were similar in magnitude for each question. Of the nine questions showing significant pre-to-post improvement in the EiE group, six questions (q3, q11R, q4, q14d, q15c, and q15g) share post-assessment scores that are not significantly different between the two genders, indicating that overall males and females perform equally on the *Designing Parachutes* assessment after receiving astronomy or solar system science instruction and then participating in the EiE *Designing Parachutes* curriculum.

Table 17. Designing Parachutes Assessment: Gender Differences (EiE only)*

	Within-Group Differences (Pre vs. Post)								Male / Female Differences	
	Female				Male				PRE	POST
Q#	N	Pre	Post	p=	N	Pre	Post	p=	p=	p=
1	582	20.6%	50.3%	**.000**	591	29.3%	58.4%	**.000**	**.001**	**.006**
7	582	57.9%	67.2%	**.001**	594	55.6%	73.1%	**.000**	.416	**.028**
11R	577	22.7%	51.6%	**.000**	588	22.3%	52.9%	**.000**	.862	.671
13R	584	30.7%	53.3%	**.000**	583	38.3%	64.5%	**.000**	**.006**	**.000**
3	583	31.0%	66.4%	**.000**	585	38.5%	71.5%	**.000**	**.008**	.061
4	587	51.3%	68.5%	**.000**	589	60.1%	71.3%	**.000**	**.002**	.291
9R	583	78.7%	80.3%	.559	586	75.8%	79.5%	.125	.227	.748
14d	585	60.5%	69.9%	**.000**	590	61.9%	74.4%	**.000**	.635	.086
15c	581	74.5%	87.4%	**.000**	590	74.2%	88.0%	**.000**	.910	.782
15g	584	87.7%	95.4%	**.000**	591	87.5%	94.9%	**.000**	.920	.718

*excluding gr6 and Asian ethnic grp from sample (because of violations of homogeneity of regression and error variance on the all_scale)

For ethnic groups, significant effects were also found after adjusting for pretest scores on the all-item scale (F(3.1519) = 15.050, p<.000, partial $\eta^{\wedge 2}$ =.029). Adjusted mean post all-item scale scores suggest that students in the White/Caucasian and “Other” ethic groups performed better on the post assessment (White mean = 6.685, “Other” mean = 6.488) than students in the Black (mean = 5.784) and Hispanic (mean = 5.972) groups. There was no significant interaction effect between the test/control groups and ethnic groups (test/control*ethnic_gps F(3,1519) = .081, p=.971, partial $\eta^{\wedge 2}$ = .000), nor was there any significant interaction between gender and ethnic groups (gender*ethnic_gps F(3,1519) = 1.083, p=.355, partial $\eta^{\wedge 2}$ = .002)

Within the EiE group, students who received free or reduced-price lunch (FRL) as well as students who did not (N-FRL) both showed significant pre-to-post improvements for nearly all questions, except that neither group improved significantly on question 9R and FRL students did not improve significantly on question 14d (Figure 14, Figure 15, and Table 18). Additionally for questions q15c and q15g (both science questions), FRL students showed pre-to-post improvements of greater magnitudes than the improvements seen by N-FRL students. However, N-FRL students performed significantly better on the post-assessment than FRL students on 8 of 10 questions; they also performed significantly better than FRL students on 5 pre-assessment questions.

Figure 14. ***Designing Parachutes*** **Assessment: Engineering Questions by FRL Status**

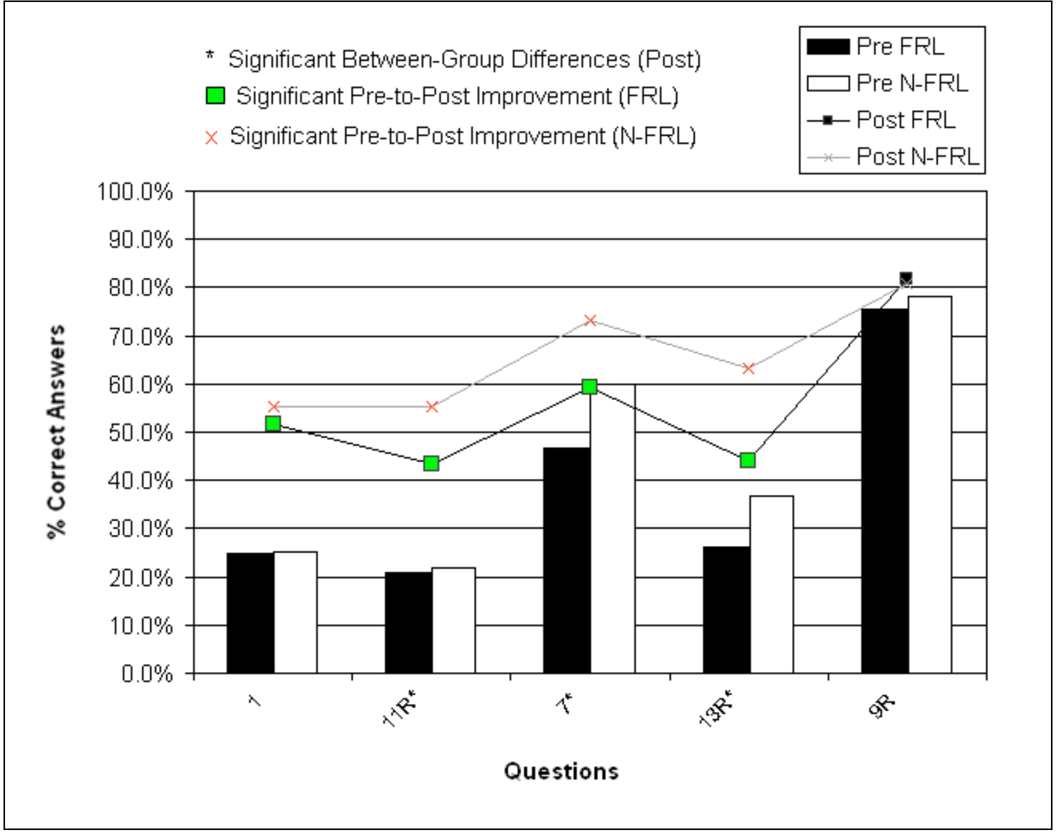
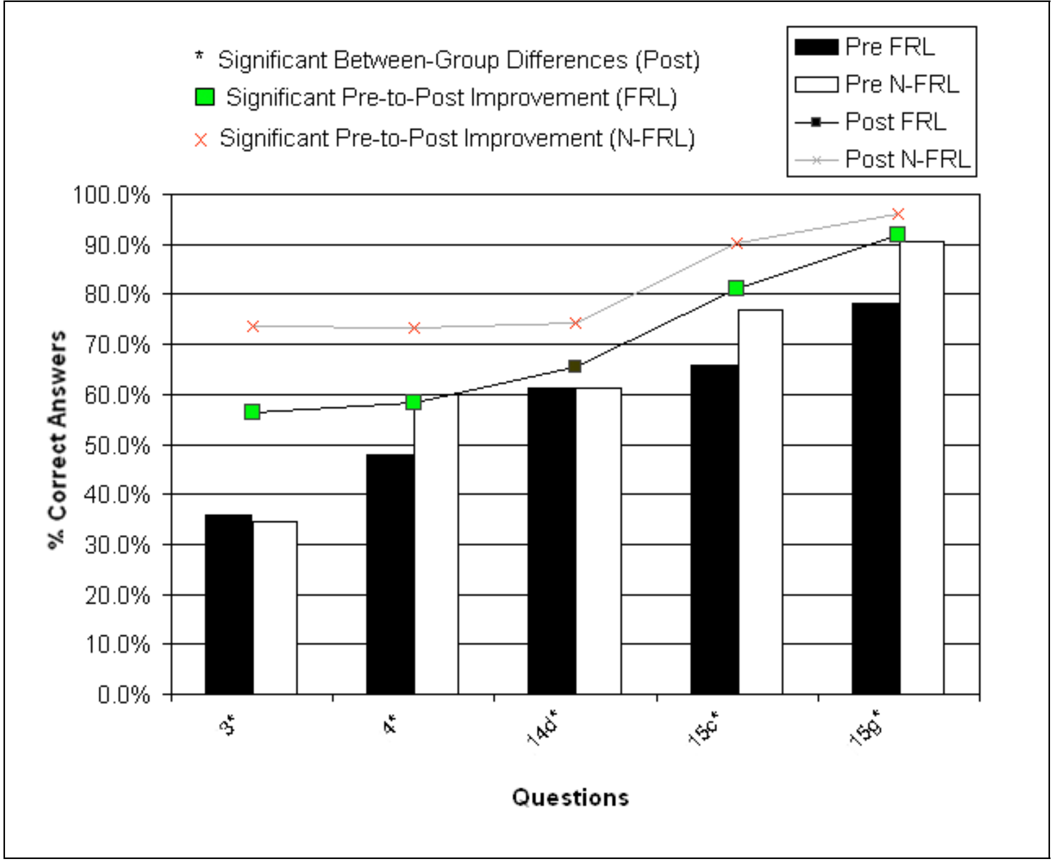


Table 18. Designing Parachutes Assessment: Free/Reduced Lunch Differences (EiE only)

	Within-Group Differences (Pre vs. Post)								FRL / Not FRL Differences	
	Free or Reduced Lunch				Not Free or Reduced Lunch				PRE-	POST
Q#	**N**	**Pre**	**Post**	**P=**	**N**	**Pre**	**Post**	**p=**	**p=**	**p=**
1	231	24.7%	51.5%	**.000**	816	25.1%	55.4%	**.000**	.890	.296
7	229	46.7%	59.4%	**.010**	821	59.9%	73.3%	**.000**	**.000**	**.000**
11R	230	20.9%	43.5%	**.000**	812	21.8%	55.3%	**.000**	.763	**.002**
13R	228	26.3%	43.9%	**.000**	813	36.9%	63.3%	**.000**	**.003**	**.000**
3	232	35.8%	56.5%	**.000**	812	34.4%	73.6%	**.000**	.689	**.000**
4	228	47.8%	58.3%	**.012**	820	59.9%	73.4%	**.000**	**.001**	**.000**
9R	225	75.6%	81.8%	.130	817	78.1%	80.8%	.198	.420	.736
14d	226	61.1%	65.5%	.387	820	61.1%	74.4%	**.000**	.992	**.008**
15c	226	65.9%	81.0%	**.000**	818	77.0%	90.3%	**.000**	**.001**	**.000**
15g	226	78.3%	92.0%	**.000**	821	90.6%	96.2%	**.000**	**.000**	**.008**

*excluding gr6 and Asian ethnic grp from sample (because of violations of homogeneity of regression and error variance on the all_scale)

Figure 15. *Designing Parachutes* Assessment: Science Questions by FRL Status



Discussion

EiE students performed better on the EiE unit assessments than control students—particularly on the engineering questions, but also on many of the science questions. Girls and boys learning from EiE both improved, as did both those who did and those who did not receive free or reduced-price lunch (FRL). Male and female EiE students tended to perform similarly, but EiE students not receiving FRL outperformed students receiving FRL. White EiE students also outperformed Black and Hispanic EiE students on both the pre- and post-assessments.

On the *Designing Solar Ovens* assessment, EiE students performed significantly better than control on the majority of post-assessment questions, including science, technology and engineering questions. Among EiE students, both girls and boys improved significantly on most questions. Girls performed significantly better than boys on two of the nine post-assessment questions. Both FRL and N-FRL students improved, though N-FRL students outperformed FRL students on both the pre-assessment and the post-assessment.

On the *Designing Parachutes* assessment, ANCOVA analysis partialling out the effect of the pre-assessment showed that EiE students performed significantly better (p<.000) on the post-assessment (adjusted mean = 6.830) than control (adjusted mean = 5.634). Looking at individual questions, EiE students performed significantly better than control on 8 of 10 post-assessment questions, most of them engineering and technology questions, but including some science questions. Among EiE students, both girls and boys improved on most questions. Boys (adj. mean = 6.425) performed significantly better (p=.016) than girls overall (adj. mean = 6.040); test sample boys performed better than test sample girls on 3 of 10 questions on the post-assessment. White students (adj. mean = 6.685) and Other (adj. mean = 6.488) ethnic groups performed significantly better (p<.000) than Black (adj. mean = 5.784) and

Hispanic (adj. mean = 5.972) students. There were no interaction effects found by test/control or by gender. Both FRL and N-FRL students improved significantly on most questions. However, N-FRL students performed better on most questions on both the pre-assessment and the post-assessment.

The low reliability scores on the *Designing Parachutes* and *Designing Solar Ovens* assessments have underscored a need for better-designed assessments. Beginning with the Year 6 pilot assessments, we have instituted a new process for developing and piloting assessments. This process includes (1) defining four target learning objectives to assess with each assessment; (2) conducting focus groups with elementary school students to better define their understanding of the target learning objectives; (3) creating at least twice as many assessment questions as it is anticipated will be needed; (4) validity testing all questions with elementary school students using think-aloud protocols; (5) conducting reliability testing with the pilot assessments to construct scales before revising the assessments into the field test versions. We anticipate that these new procedures will result in better data for analysis next year.

Due to the inadequacy of the assessments and the irregularity of data collection—with some teachers participating in both control and test, and others participating in one or another, but overall fewer control than test classrooms—these results should be considered promising, but not final. A fully funded, properly controlled study is called for to confirm the findings of this evaluation.

References

---

[1] Cunningham, C.M. and K. Hester, 2007, "Engineering is Elementary: An Engineering and Technology Curriculum for Children", presented at American Society for Engineering Education Annual Conference & Exposition, Honolulu, HI.

[2] Bonate, Peter L., 2000, *Analysis of Pretest-Posttest Designs,* Boca Raton, FL: Chapman & Hall/CRC.