The Boston Museum of Science BEST Project

Annual Evaluation Report

Submitted by Russell Faux, Ed.D.
February 27, 2012

davisSquare

Executive Summary

The following report presents the findings from the evaluation research conducted by Davis Square Research Associates (DSRA) for the Boston Museum of Science Bridging Engineering, Science and Technology (BEST) project, funded by the National Science Foundation, and designed to strengthen the teaching and learning of engineering content in elementary school teacher preparation programs. The research focus is on the measurement of the extent to which participating college students' judgments regarding engineering came to resemble the judgments of engineering experts, as well as other changes around attitudes toward engineering that can be associated with participation in BEST. A secondary area of study is the effect of attending a workshop for college faculty, held at the Museum of Science in June 2011.

Key findings include:

- Fall 2011 participant judgments on statements related to engineering changed to resemble more closely the judgments on the same statements as made by engineering experts, though spring 2011 students' judgments showed little change.
- Student attitudes toward engineering showed sizable improvements from pre- to post-test.
- Preservice teacher attitudes also showed strong improvements toward the place of engineering teaching and learning in their future classrooms.
- Faculty members who attended the June 2011 workshop reported very strong and positive gains in their attitudes toward teaching engineering, though the integration of engineering into other courses was reported as a major challenge.

Sample & Method

There are two student samples for the current report, with the first coming from the spring 2011 semester, while the second sample comes from the fall 2011 semester. All students were exposed to Museum of Science Boston Engineering is Elementary content and resources as a part of the regular curriculum. The comparison group for the study comprises 94 engineers and professors of engineering from across the United States. The sample for the workshop includes 24 faculty members all of whom have made a commitment to incorporating EiE content in their teaching.

The following two tables summarize the treatment samples for this report.

Table 1: Spring 2011 Students by College

	Pre-Test	Post-Test	Total
Berkshire CC	55	28	83
Fitchburg State	6	56	62
MCLA	26	10	36
Massasoit CC	52	77	129
Middlesex CC	70	47	117
Salem State University	68	108	176
Total	277	326	603

Table 2:Fall 2011 Students by College

	Pre-Test	Post-Test	Total
Berkshire CC	36	0	36
Bridgewater State	13	12	25
Fitchburg State	126	65	191
Massasoit CC	35	0	35
Middlesex CC	47	22	69
North Shore CC	55	44	99
Salem State University	6	26	32
Total	413	263	676

Data from the implementation of the units are lacking in this report, and, as such, the findings ought not to be interpreted as being as definitive as they might sound. There are many variables that can come into play in any given implementation, and given that multiple classes from eight Massachusetts colleges were involved in this round of the research, the likelihood of there being considerable variation is very high.

The data from the survey items related to the participants' skills at judging the correctness of a series of 31 statements about engineering. These statements (presented below in the item-by-item presentation) have already been judged by a group of 94 experts in engineering and engineering education from across the US. The point of the instrument is to see whether the judgments expressed by the student participants come to more closely approximate those of the experts.

On the post-test survey, respondents were asked a series of retrospective pre-test questions about their knowledge of, and attitudes toward engineering. In addition, the post-test survey included a series of retrospective pre-test questions for preservice teachers regarding their expectations for using engineering in their work with children.

The key questions for the survey, as first developed in the 2011 report, continue to be:

- Have the students' responses to ambiguous statements about engineering come closer to those of engineering professionals?
- What changes do respondents report with regard to their knowledge of, and attitudes toward engineering?
- How do preservice teachers envision the role of engineering in their teaching?
- What are the effects of the professional development workshop conducted in June 2011?

Survey Findings

This section examines first the gains that the students may have made in the responses to the judgments survey, and then DSRA presents the preliminary findings from a validation study of the instrument itself.

Engineering Judgments Survey

This section presents the findings first from the spring 2011 semester and then from the fall 2011 semester.

The following table summarizes the instances of significant difference between the spring 2011 participants and the engineering judgments expressed by the expert group. The 31 items in the survey are divided into three domains (previously validated through a principal components analysis). The “process” items are by far the largest set of items, with these statements referring to various aspects of the engineering design process (a central feature of the EiE program). The “context” items refer to the circumstances surrounding the work of engineers, and the “products” refer to those items that refer to the tangible results of engineering activities. Note that there was very little movement from pre- to post-test among the spring 2011 students, while the fall students showed a far greater improvement.

Table 3: Spring 2011 Pre-Post Changes Summarized

College	Process (N=19)		Context (N=5)		Products (N=7)	
	Pre-	Post-	Pre-	Post-	Pre-	Post-
Berkshire CC	16	15	5	5	7	7
Fitchburg State	18	18	5	5	7	7
MCLA	17	17	5	5	7	7
Massasoit CC	17	17	5	5	7	7
Middlesex CC	18	18	5	5	7	7
Salem State University	17	17	5	5	7	7
Total	103/114	102/114	30/30	30/30	42/42	42/42
Percentage of Items Showing Significant Difference	90%	90%	100%	100%	100%	100%

The reliability value for the survey was a very good 0.79 (Cronbach Alpha), with all items showing a greater-than-expected consensus (Kolmogorov-Smirnov), meaning that the students tended to vary relatively little on their responses to the survey items. Overall, DSRA found 175 instances of pre-test differences between the students and experts (of 186 possible instances of difference, or nearly 94%), while at the post-test DSRA found 174 (of 186 possible instances, or again about 94%). The pre-post change here is next to imperceptible, with the students performing at almost exactly the same level both before and after exposure to the unit. DSRA hastens to underscore that this does not mean the students did not improve at all, but simply that the judgments instrument could not detect any changes.

The fall 2011 students acquitted themselves much better on the judgments assessment, with the gap between experts and students narrowing consistently and by a considerable margin.

Table4: Fall 2011 Pre-Post Changes Summarized

College	Process (N=19)		Context (N=5)		Products (N=7)	
	Pre-	Post-	Pre-	Post-	Pre-	Post-
Fitchburg	18	8	5	3	7	2
Massasoit	18	-	5	-	7	-
Middlesex CC	17	5	4	3	7	1
Salem State	12	14	1	5	7	5
Berkshire	16	-	5	-	7	-
North Shore CC	17	17	5	5	7	6
Bristol CC	16	2	5	2	7	4
Bridgewater	18	9	4	3	7	2
Total	132/152	55/114	34/40	21/30	56/56	20/42
Percentage of Items Showing Significant Difference	87%	48%	85%	70%	100%	48%

For this second group, DSRA found 222 instances of pre-test differences (of 248 possible instances of difference, or nearly 90%), while at the post-test DSRA found 96 (of 186 possible instances, or about 52%). This is an enormous gain, with the differences between the students' judgments and those of engineering experts strikingly reduced. Of particular note is the reduction in the expert/student differences in the domains of process and products. The reduction in the domain of context is noticeably smaller, suggesting that the implementations across the colleges tended to underemphasize the contexts of engineering in society. The reliability value for the survey was a good 0.73 (Cronbach Alpha), with only five pre-test items and 16 post-test items showing a greater-than-expected consensus (Kolmogorov-Smirnov), meaning that the students varied less on their responses to the survey items on the post-test and they had on the pre-test.

Attitudinal Changes

In this section DSRA summarizes the retrospective pre-post data from the second section of the post-survey. There are eleven pairs of questions that are intended to gather information on changes on the respondents' attitudes toward engineering. All eleven pairs showed significant improvement (Wilcoxon), with small changes in the standard deviation for both the spring and fall samples. The overall pre-post change was also significant, with an effect size of 0.61 for the spring semester students and 0.43 (both often characterized as a "moderate" magnitude). This is a very solid finding, though the gains were found to be normally distributed, meaning that some reported greater gains than did others. The pre-test was significantly correlated (Pearson, 0.61 for the spring students and 0.35 for the fall 2011 students), meaning that students who had lower values at the pre-test had lower values at the post-test, and vice-versa. The scale used was a Likert-type scale running from 1 (very low/negative) to 6 (very high/positive). Note that the spring group reported generally lower attitudes at both the pre- and post-test, though this same group made greater overall gains.

Table5: Reported Pre-Post Attitudinal Changes by Item

Item	Mean: Spring 2011	Mean: Fall 2011
I like to learn about engineering - BEFORE	2.98	3.04
I like to learn about engineering - AFTER	4.33	4.40
I like to think about engineering - BEFORE	2.66	2.78
I like to think about engineering - AFTER	3.88	3.96
I notice engineering in the news - BEFORE	2.64	2.88
I notice engineering in the news - AFTER	3.85	3.92
I like to talk about engineering - BEFORE	2.14	2.36
I like to talk about engineering - AFTER	3.14	3.46
I feel like I can learn about engineering - BEFORE	3.23	3.39
I feel like I can learn about engineering - AFTER	4.50	4.56
I feel like I can be good at engineering - BEFORE	2.56	2.57
I feel like I can be good at engineering - AFTER	3.68	3.75
I feel like I will benefit from learning more about engineering - BEFORE	3.12	3.22
I feel like I will benefit from learning more about engineering - AFTER	4.52	4.48
I admire people who are in engineering - BEFORE	3.77	3.92
I admire people who are in engineering - AFTER	4.88	4.85
I think about pursuing further study in engineering - BEFORE	2.00	1.64
I think about pursuing further study in engineering - AFTER	2.65	2.34
I think about how engineering will fit into my career - BEFORE	2.13	2.19
I think about how engineering will fit into my career - AFTER	3.18	3.51
I pay attention to the place of engineering in society - BEFORE	2.42	2.52
I pay attention to the place of engineering in society - AFTER	3.94	3.91

Attitudes toward Engineering – Pre (max=66)	29.70	34.40
Attitudes toward Engineering – Post (max=66)	42.62	48.45

Preservice Teacher Survey

The final section of the online survey was for preservice teachers only, and so the sampling frame is somewhat smaller, with the spring group (N=99) slightly smaller than the fall group (N=105).

The following table summarizes the findings for the seven paired items intended to gather information on the respondents' attitudes toward the place in engineering in the respondents' future teaching. Each pair was found to show significant positive change (Wilcoxon), as was the overall change for both groups. The magnitude of the change was 0.56 for the spring 2011 group and 0.52 (eta-squared, a moderate size change). The gains for both groups were normally distributed (Kolmogorov-Smirnov) meaning that some respondents reported greater gains than did other. In general, students who reported lower pre-test values tended to report lower post-test values, with the pre-post scores significantly correlated (Pearson, 0.53 for the spring group and 0.52 for the fall group). As with the previous section, the reported degree of participation in the lesson was not a significant predictor of overall gains for either group.

The scale used for the following items was a Likert-type scale running from 1 (very low/negative) to 6 (very high/positive).

Table 6: Pre-Post Preservice Teacher Attitudinal Change by Item

Item	Mean: Spring 2011	Mean: Fall 2011
I feel like I can help children to learn about engineering - BEFORE	2.45	2.10
I feel like I can help children to learn about engineering - AFTER	3.98	3.88
I feel like I can guide children through a class engineering project - BEFORE	2.51	2.32
I feel like I can guide children through a class engineering project - AFTER	4.05	4.01
I think my students will have a good awareness of engineers - BEFORE	2.51	2.37
I think my students will have a good awareness of engineers - AFTER	3.91	3.77
I think my students will have little knowledge of engineering - BEFORE	2.95	2.72
I think my students will have little knowledge of engineering - AFTER	3.50	3.19
I think my students will not be able to learn the vocabulary they will need to do class engineering projects - BEFORE	2.72	2.50
I think my students will not be able to learn the vocabulary they will need to do class engineering projects - AFTER	3.72	3.15
I think my students will enjoy doing class engineering projects - BEFORE	3.14	3.11
I think my students will enjoy doing class engineering projects - AFTER	4.59	4.44
I think my students will learn a lot from doing class engineering projects - BEFORE	3.30	3.35
I think my students will learn a lot from doing class engineering projects - AFTER	4.69	4.60
Attitudes toward Teaching Engineering – Pre (max=42)	19.76	18.48
Attitudes toward Teaching Engineering – Post (max=42)	28.53	27.31

June 2011 Participant Workshop

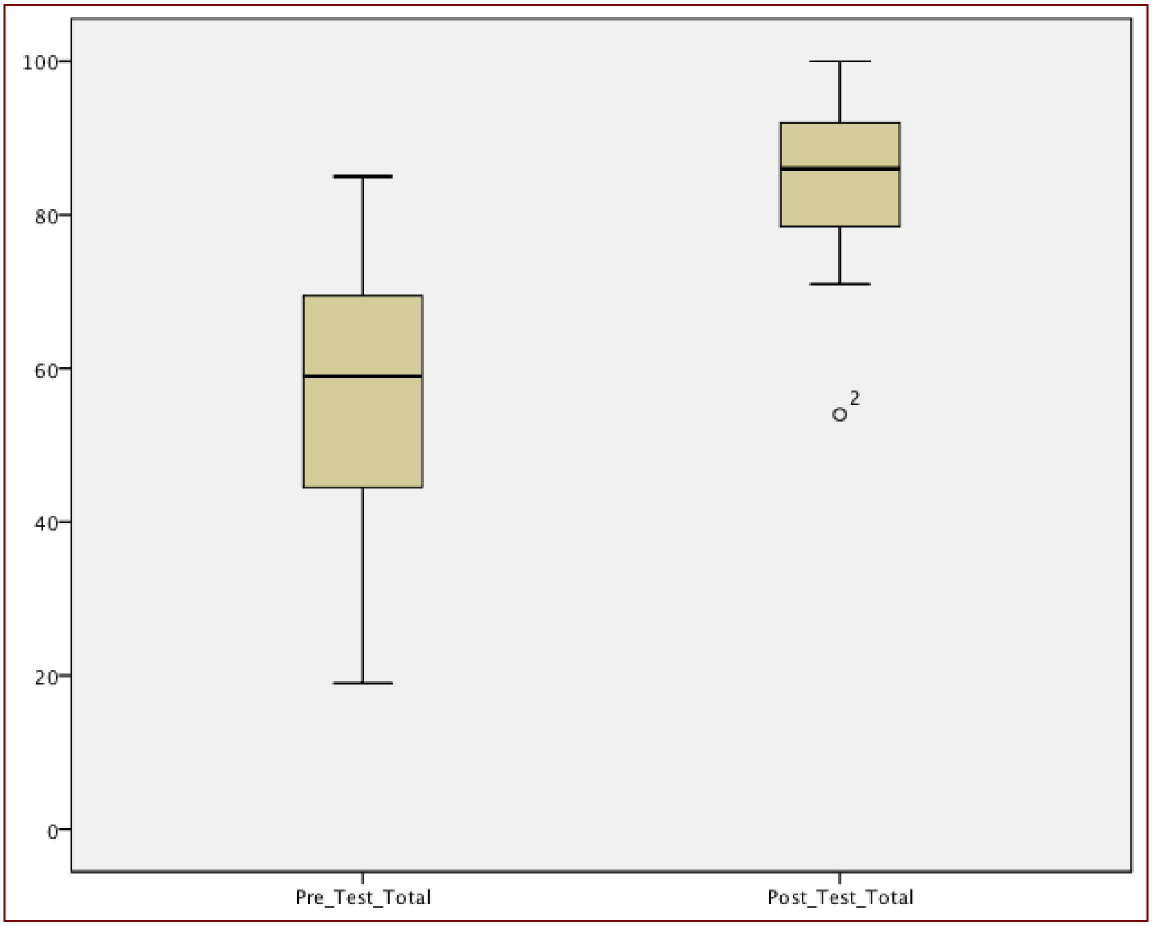
At the conclusion of the June 2011workshop for faculty who will be incorporating EiE content into their courses, all participants completed an online survey intended to gather information on the effects of the activities. The survey used 16 retrospective pre-test questions ("before the workshop" and "after the workshop") to gather data on the changes in the attitudes of the attendees (N=24). The instrument included items that focused on respondents' attitudes toward engineering, toward their own skills at teaching engineering, and toward their expectations for student learning. The reliability value for the survey was a very high 0.96 (Cronbach Alpha), with all items showing significant (Wilcoxon) pre-post change. The overall effect size (eta-squared) was sizeable 0.75. However, gains were normally distributed, meaning that some people gained more and others gained less. Note that for the items in the following table, participants moved from 59% of the maximum to a full 87%. This increase in attitudinal responses is a very good finding and one that is essential for longer-term project effectiveness.

Table 7: Reported Pre-Post Change Among Workshop Attendees

Item	**Mean: Pre**	**Mean: Post**
Enthusiasm for teaching engineering	4.22	5.43
Interest in learning more about engineering	4.26	5.35
Your skills at teaching the engineering design process	4.04	5.22
Comfort in using hypotheses in teaching engineering	4.43	5.22
Your ability to present engineering concepts in engaging ways	3.78	5.17
Your confidence that students will enjoy learning about engineering	4.00	5.13
Confidence in designing student engineering explorations	3.70	5.04
Your ability to spark student interest in engineering	3.78	5.04
Your ability to draw on everyday examples to explain design judgments	3.91	5.00
Your skills at facilitating engineering explorations	3.61	4.87
Knowledge of engineering for teaching	3.65	4.87
Your expectations for student learning in engineering	3.48	4.83
Your expectations for student interest in engineering	3.74	4.61
Your expectations for student attitudes toward engineering	3.35	4.48
Your ability to understand student misconceptions in engineering	3.30	4.43
Your ability to anticipate student misconceptions in engineering	3.17	4.35
Total (maximum=96)	56.70	83.87

The following boxplots illustrate the pre-post change (the small 2 refers to participant #2, an outlier).

Table 8: Pre-Post Distributions



When asked about the impediments to a larger implementation of engineering, the respondents consistently cited the "crowded curriculum," with this item being the only item that achieved statistical significance (Kolmogorov-Smirnov). Reponses varied normally to all other items in this forced ranking set of questions. This finding seems to indicate that the participants are still wrestling with the integration of engineering into pre-existing content, though the data on this point are not complete.

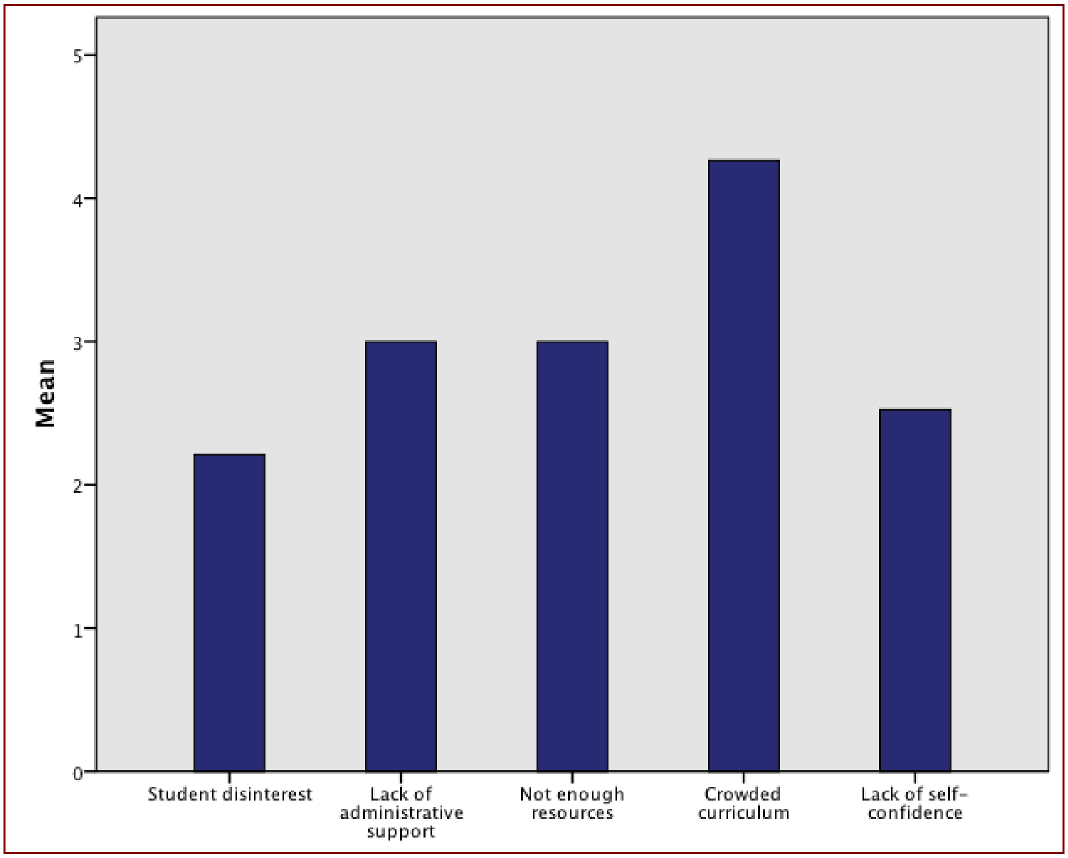
Table 9: Reported Impediments to a Wider Implementation

Impediments	Mean	Std. Deviation	Modal Response	% Modal
Crowded curriculum*	4.26	1.10	5	65%
Lack of administrative support	3.05	1.00	2	36%
Not enough resources	2.86	1.39	4	33%
Lack of self-confidence	2.41	1.37	1	36%
Student disinterest	2.29	1.45	1	39%

*Significant at $p<.05$ (Kolmogorov-Smirnov)

The following bar graph presents the distribution of the reported impediments.The tendency to cite the crowded curriculum is readily seen in this graphic. All other impediments were reported as varying in importance by workshop participants.

Figure 1: Impediments to a Wider Implementation



Respondents were also asked about ways in which participation in the project contributed to their improvement as a teacher. All 24 responses were highly positive, with these being very roughly divisible into responses that seem to index personal growth and those that appeared linked to improved classroom performance. The following table offers several examples from each category. Note that this division is hardly watertight, but rather organizes the responses into general trends that likely have considerable overlap.

Table 10: Examples of Effects of Participation

Personal	Professional
• To my surprise, I have found that I have a natural aptitude for engineering. My father and brothers all have this, but I thought it had passed me by. Recognizing my own capacity for engineering gives me the confidence to teach it. • I feel energized! The networking and exposure to new ideas has energized me to develop new materials for teaching. • I feel confident teaching the EiE Engineering process • I feel confident in leading investigations in engineering. My content knowledge is in life science so I am learning about other areas but still need more confidence in math and physics.	• It has helped me see that I can embed math curriculum into engaging engineering activities • I have found an exciting way to help non-majors understand how engineering and technology is involved in their every day lives and how closely related it is to the other sciences especially biology. • It has demonstrated an incremental/scaffolding approach with engaging material as a path to more student engagement. I believe more so than ever that sacrificing content for deeper learning is a better path for my student population. • I gained wonderful ideas about incorporating the simplified engineering design process into activities for lab. • Providing hands-on experiences I can use in the classroom.

Conclusions & Recommendations

The above data show that the spring 2011 students are rather anomalous in their lack of progress as measured by the judgments survey. Both the previous cohorts, as well as the following cohort showed very strong gains on the instrument. The reason for this "dip" in the performance cannot be ascertained on the basis of the available data, and speculation here would be incautious. Both cohorts did very well on the attitudinal surveys, with the two groups reporting highly comparable gains. Results varied little from school to school, suggesting generally similar implementations, though again, little is known about the implementations.

The pre-service teacher attitudinal survey items showed strong improvements in respondents' attitudes toward the content of engineering, their own skills at teaching engineering to young children, and their intentions to use engineering in their work. Such findings are very encouraging, and they indicate that the pre-service teachers feel inspired and capable of including engineering. Similarly, the June 2011 faculty workshop generated strongly positive attitudinal data, this time with a somewhat larger effect size.

Taking into consideration the above, DSRA recommends the following:

- That the project develop a fidelity of implementation rubric that can be included in the online surveys of participants. Such a rubric should collect data on participant characteristics, resources utilized, and strategies employed.
- That the project consider looking at longer term effects of participation, especially among the preservice teachers, for the purpose of measuring any enduring effects of participation.
- That the project considering publishing a report that explores (1) the relationship between EiE content and college-level engineering and (2) the unique educational contributions of the EiE content.