MIAMI UNIVERSITY
DISCOVERY CENTER
Evaluation • Research • Professional Learning

Evaluation of CEEMS: The Cincinnati Engineering Enhanced Mathematics and Science Partnership Project

ANNUAL REPORT 2017-2018

DISCOVERY CENTER
for
EVALUATION, RESEARCH, AND PROFESSIONAL LEARNING
Formerly Ohio's Evaluation & Assessment Center

MIAMI UNIVERSITY
OXFORD, OH

Please cite as follows:

Dixon, M. L., Li, Y., & Woodruff, S. B. (2018). *Evaluation of CEEMS: The Cincinnati Engineering Enhanced Mathematics and Science Partnership Project: Annual report, 2017-2018.* Oxford, OH: Miami University, Discovery Center for Evaluation, Research, and Professional Learning.

Distributed by:

408 McGuffey Hall
210 E. Spring St.
Oxford, Ohio 45056
discoverycenter@MiamiOH.edu
(513) 529-1686 phone
(513) 529-2110 fax

Table of Contents

Table of Tables

Table of Figures

Executive Summary

Miami University's Discovery Center for Evaluation, Research, and Professional Learning (Discovery Center) completed the final year of a 3-year mixed methods evaluation of instructional change for teachers involved with the *CEEMS: The Cincinnati Engineering Enhanced Mathematics and Science* partnership project. The CEEMS partnership was a multi-year, multi-million dollar NSF-funded initiative. The purpose was to create multiple professional development opportunities and teaching certification pathways for pre-service and in-service K-12 teachers to develop skills necessary to teach engineering as part of the science curriculum.

In 2015, CEEMS contracted with the Discovery Center to evaluate the ways teachers and Resource Team members changed during the course of their participation in CEEMS. This evaluation asked the following questions:

1. In what ways did teachers' instructional practices change in the course of their participation in CEEMS?
2. In what ways did CEEMS Resource Team support for teachers change in the course of their participation in CEEMS?

The Discovery Center collected qualitative focus group and both qualitative and quantitative classroom observation data to respond to these evaluation questions. This report provides a summative assessment of teacher and Resource Team member change associated with CEEMS.

In what ways did teachers' instructional practices change in the course of their participation in CEEMS?

Teachers used a wider variety of instructional practices in CEEMS lessons than in traditional lessons. Instructional practices evident in CEEMS lessons included probative, open-ended questioning that encouraged critical thinking; the engineering design process; challenge-based learning strategies; collaborative grouping; and external resources (e.g., videos) as a means to focus the lesson on real-world issues.

Teachers supported their instructional practices by drawing upon a network of other CEEMS teachers and Resource Team members that emerged as a direct result of the CEEMS program. This network generated a vast amount of curriculum resources, such as lesson plans, that were shared and made available to the wider public electronically.

Change in instructional practice was dependent upon individual effort and inclination toward change and was not guaranteed to result from participation.

In what ways did CEEMS Resource Team support for teachers change in the course of their participation in CEEMS?

The Resource Team developed and improved upon a set of resources—such as instructional guides—and processes—such as post-unit debriefing—through trial and error over the years of the program. These improvements enhanced the teachers' experiences by facilitating and streamlining communication and providing a wealth of resources that had been applied in real classroom settings. The CEEMS project team's willingness to incorporate suggestions and changes initiated by the Resource Team enabled these improvements.

The Resource Team developed a collaborative coaching structure in which multiple Team members had multiple contacts with individual teachers over time. The collaborative structure allowed Resource Team members to learn from one another and draw upon a common set of resources.

The Resource Team developed a number of coaching strategies to support instruction that was responsive to a challenge-based learning environment. Some of the most important coaching strategies supported teachers' ability to develop new challenge-based engineering design units and lessons consistent with grade-level standards, ask probing questions that sparked critical thinking, ensure science/mathematics content learning occurred alongside engineering design activities, manage a classroom environment in which students worked in small groups on hands-on activities, and use challenges to connect with building administrators in positive ways.

Individual Resource Team members who were former engineers discussed engineering as a career within classrooms, thus supporting student understanding of engineering in the real world.

Introduction

The Discovery Center for Evaluation, Research, and Professional Learning (Discovery Center) completed the third year of a 3-year mixed-methods evaluation of instructional change for teachers involved with *CEEMS: The Cincinnati Engineering Enhanced Mathematics and Science* partnership project. The broad objective of the evaluation was to provide annual feedback and a final summative assessment of the project's ability to meet revised project Goal 3, "Develop math and science teacher knowledge of challenge-based learning, engineering, and the engineering design process as instructional strategies through explicit training and classroom implementation support" (Maltbie & Butcher, 2014).

The purpose of this study was to evaluate the influence of CEEMS participation on teachers' confidence and competence in their incorporation of engineering design principles into their science instruction. Specifically, this evaluation asked the following questions:

1. In what ways did teachers' instructional practices change in the course of their participation in CEEMS?
2. In what ways did CEEMS Resource Team support for teachers change in the course of their participation in CEEMS?

The CEEMS partnership project was funded through a Math and Science Partnership (MSP) program grant from the National Science Foundation (NSF). This project was a multi-year effort to create multiple professional development pathways to prepare pre-service and in-service teachers to meet revised standards for engineering education in the K-12 science curriculum.

For this project, the Discovery Center Evaluation Team consisted of Dr. Sarah B. Woodruff, Principal Investigator for the evaluation; Ms. Yue Li, Senior Research Associate and Project Team Leader; and Ms. Maressa Dixon, Research Associate. This annual report describes evaluation activities conducted in the 2017-2018 academic year, findings from those activities, and triangulated findings from all three years of the evaluation. Descriptions of evaluation activities for the first two years are available in evaluation reports from those years (Woodruff, Dixon, and Li, 2016; Woodruff, Dixon, and Li, 2017).

Table 1 summarizes alignment between evaluation questions and evaluation measures.

Table 1. *Alignment of Evaluation Questions and Measures*

Evaluation Question	Instrument/ Measure(s)
EQ 1: In what ways did teachers' instructional practices change in the course of their participation in CEEMS?	Classroom Observations Principal Focus Group Resource Team Focus Groups
EQ2: In what ways did CEEMS Resource Team support for teachers change in the course of their participation in CEEMS?	Resource Team Focus Groups

In the 2017-2018 academic year, the Discovery Center Evaluation Team conducted one focus group with CEEMS Resource Team members and analyzed 97 classroom observations from 15 CEEMS teachers (cohort 5). Additionally, the evaluation team trained 5 CEEMS engineering fellows to conduct classroom observations. The next section of this report describes evaluation methods and findings.

Evaluation Methods and Findings

Resource Team Focus Group, Fall 2017

The Discovery Center invited CEEMS Resource Team members to participate in a focus group held on February 7, 2018; five Resource Team members participated.

Instrument

The Resource Team focus group protocol included questions about the ways Resource Team members supported teacher instruction and CEEMS program participation. Specifically, the focus group protocol asked about: differences in the coaching role in the final year of the grant; how Resource Team members maintained quality control in the production of deliverables; how Resource Team members were limited in their ability to support teachers; and how Team members supported teachers in terms of classroom management, time management, content learning, and overcoming resistance. Participants were asked to offer additional information that was not asked about at the end of the focus group. A copy of the focus group protocol is available in Appendix A.

Data Collection

The Research Associate on the Evaluation Team conducted the focus group, and the same Evaluation Team member analyzed the data with the support of a graduate research assistant. Prior to beginning the focus group, participants granted the evaluation team informed consent to record the conversation and use the data to evaluate the CEEMS program. The focus group lasted about an hour, and the conversation was recorded using a digital hand-held recorder. Digital audio data were uploaded to the data analysis software package NVivo, version 11, for transcription. A graduate research assistant created "clean" transcripts—transcripts that are close to verbatim but do not include pauses, misspoken words, or extraneous utterances—of the focus group.

Data Analysis

To analyze these data, the Evaluation Team member read participant responses, question by question, to identify themes related to each protocol question and salient themes across questions. The analysis approach was inductive, in that responses were coded based on the nature of the response. Themes were declarative statements that summarized the data in ways that responded to the protocol questions and the overall evaluation questions.

Findings

Findings organized by protocol question were provided to the CEEEMS Project Team in April 2018 as part of an evaluation brief. Findings presented in this report were synthesized to respond directly to the evaluation questions.

Patterns of Teacher Change

By the 2017-2018 academic year, all participating CEEMS teachers had been a part of the program for at least one year and a network of support from the Resource Team and former CEEMS teachers had been built for more than five years. As a result, cohort 5 teachers knew what to expect from the CEEMS program and from their students and had access to resources that previous cohorts had already developed and applied. For teachers whose change was demonstrable to the Resource Team, the nature of change responded to the needs of students in a challenge-based (as opposed to traditional) instructional environment. These changes most notably occurred in the areas of classroom management, content instruction, connecting content and challenge activities, and valuing the challenge-based approach as part of the science or mathematics classroom.

Classroom management. Some teachers developed classroom management skills appropriate to challenge-based learning as a result of CEEMS participation. CEEMS teachers and the Resource Team learned, for example, that providing time for instructions and questions at the beginning and debriefing at the end of the lesson facilitated learning in challenge lessons much better than interrupting student activities in the middle of the challenge to provide instruction. For some teachers, improvements to classroom management occurred after observing veteran teachers and/or having Resource Team members who were former educators observe their lessons and provide feedback.

Content instruction. Teachers and the Resource Team learned that separating instruction on content from the challenge aspects of the unit helped ensure the unit progressed in the expected time frame. Teachers allowed content-focused lessons to unfold at the pace of student learning and focused CEEMS units narrowly on the challenge and engineering design activities. As a result, teachers in cohort 5 were better able to predict the length of CEEMS units than teachers in earlier cohorts.

Connecting content and challenge activities. Teachers in later cohorts were better able to embed content into the challenge or design portions of CEEMS units through the use of forms and processes to document different aspects of the content than were earlier teacher cohorts. For example, teachers developed and shared forms for students to create an initial sketch of a design, then make a second sketch that documented and justified modifications from the first sketch. These modifications were intended to be informed by student understanding of the science or mathematical concept taught in earlier lessons. In contrast, teachers in early cohorts tended to allow students to make design changes based on what they believed would improve the design, rather than scientific or mathematical principles.

Valuing the challenge-based approach. Many teachers were skeptical of the extent to which CEEMS could work in their particular classrooms until after they saw it work well. This acceptance after initial skepticism was particularly evident among teachers who initially believed their students would not respond in expected ways to the challenge-based approach to learning. These teachers became convinced of the value of the CEEMS approach after they witnessed a dramatic change in student enthusiasm, engagement, and/or learning during challenge-based CEEMS units.

Patterns of Resource Team Change

Similar to CEEMS teachers, the Resource Team improved their practice as a result of experience, trial-and-error, the development of a CEEMS network, and improved access to resources. For the Resource Team, the nature of change responded to the needs of teachers as professionals and as learners. These changes most notably occurred in the areas of relationship building, communication, the sharing of resources, and the development of questioning techniques that facilitated student learning in challenge-based instructional environments.

Relationship building. Resource Team members learned to walk the thin line between offering gentle suggestions and telling teachers what to do, as the nature of communication affected the nature of the relationship between the Resource Team member and the individual teacher. Developing a supportive, peer-to-peer, non-authoritative relationship was crucial for the Resource Team to embody the dual roles of providing instructional coaching and accountability to CEEMS requirements.

Communication. The Resource Team learned that building relationships with individuals helped them determine how best to communicate with teachers and with one another to support teachers' needs. As the CEEMS program progressed, Resource Team members began to work in pairs to provide overlapping (i.e., initial and follow-up) communication with teachers and maintain structure during implementation of CEEMS units. Resource Team members also responded to teachers' needs by communicating via text, e-mail, and video conference—rather than solely in person—based on a teacher's individual preferences.

Sharing of resources. The Resource Team facilitated networking among teachers as well as resource development and sharing. For example, Resource Team members connected inexperienced teachers with veteran teachers when inexperienced teachers were likely to benefit from observing veteran teachers' classes. Resource Team members often shared resources developed by one teacher with the entire CEEMS network through the CEEMS website and through the Summer Institute for Teachers (SIT) professional learning event that occurred every summer. Additionally, the Resource Team shared resources with the CEEMS network that they either developed on their own or identified elsewhere, often in response to a teacher need they observed as part of regular instructional coaching duties.

Development of questioning techniques. One important instructional coaching role the Resource Team played was to support teachers' development of skills relevant to challenge-based learning that they may not have developed in traditional classroom settings. One skill the

Resource Team cultivated among teachers was the ability to question students in ways that elicited critical thinking rather than recitation of memorized facts. Part of that skill included the ability to elicit questions from students in ways that supported student conceptual understanding. Questioning techniques that facilitated critical thinking included requiring students to choose one team question when students worked on a challenge in small groups, pausing between questions to enable student thinking before providing a response, and ensuring students that there were no right or wrong answers during discussions of their thoughts and opinions.

Summary

During the course of participation in CEEMS, Resource Team change often reinforced teacher change toward instructional practices that reflected challenge-based learning. These changes occurred through trial-and-error and the accumulation of knowledge and resources through experience. Teachers developed classroom management skills that reflected the needs of students in challenge-based learning environments, separated content instruction from engineering design to streamline CEEMS unit implementation without short-changing student content learning, developed methods to assess content learning during design experiences, and valued challenge-based instruction as a result of participation in CEEMS. The Resource Team's experiences in CEEMS classrooms early in the program were instrumental in supporting those changes. For the Resource Team, change included the development of peer-to-peer relationships with teachers, the use of multiple communication strategies based on teacher preference, the sharing of resources to provide a network of support and a base from which teachers could build, and the cultivation of questioning strategies appropriate to the challenge-based instructional environment among teachers. Although teacher change was individual to the teacher and not all teachers followed the same pattern of change, the Resource Team provided an invaluable combination of instructional support and accountability that enhanced the implementation and value of the CEEMS program.

Classroom Observations, 2017-2018 Academic Year

Instrument

To understand change in teachers' instructional practices, classroom observation data were collected through hand-written field notes and completion of the *Inside the Classroom Observation and Analytic Protocol* (Horizon Research Inc., 2000). This mixed methods protocol included sections that allowed observers to record basic information about the classroom—such as the subject and number of students—and then provide both quantitative ratings and narrative descriptions of important elements of the lesson. These elements included lesson purpose, focus, design, implementation, and content; classroom culture; time usage; likely impact on mathematics/science learning; specific lesson features; and overall lesson quality. Observers also provided a narrative description of what occurred during the observation. A copy of the observation instrument can be found in Appendix B.

Data Collection

In Fall 2017, the Research Associate on the Evaluation Team trained 4 CEEMS engineering fellows to conduct classroom observations using the *Inside the Classroom Observation and Analytic Protocol*. One engineering fellow was replaced by a new fellow, who was trained in Spring 2018. The training sessions were identical and involved an introduction to low-inference observation techniques, a detailed review of the protocol, discussion of the meaning of protocol items, and practice with the instrument using video-taped lessons.

Data collection was completed by these trained engineering fellows, and fellows provided completed observation protocols to the Evaluation Team for analysis. Fellows observed 97 lessons from 15 CEEMS teachers (cohort 5) in the 2017-2018 academic year. Fellows observed four types of lessons, i.e., non-CEEMS, CEEMS Introductory, CEEMS Challenge, and Comparison lessons. Non-CEEMS lessons focused on the development of the conceptual understanding necessary to complete CEEMS challenges; these lessons were part of a CEEMS unit, but were not intended to introduce or implement the CEEMS challenge. In CEEMS Introductory lessons, the teacher introduced a challenge students would work on for several lessons. In CEEMS Challenge lessons, the teacher engaged students in one or more stages of CEEMS challenge implementation. Comparison lessons were lessons from a different unit unrelated to the CEEMS unit. Classroom observation data were entered into an Excel spreadsheet to facilitate analysis.

Quantitative Data Analysis

The Evaluation Team designed analyses to take into account unevenness in the number of observations by lesson type for each teacher. First, the Evaluation Team used descriptive statistics (e.g., frequencies; percentages), one-way ANOVA, and Pearson's chi-square tests to understand lesson features and similarities and differences in lesson ratings, by lesson type, for the full sample. Next, the Evaluation Team identified teachers with at least one observation for three different lesson types (n=10). These data were analyzed using repeated-measures one-way ANOVA, with significant findings followed by Friedman Tests and paired-samples t-tests to identify which pair of lesson types were statistically significantly different. This sub-sample was identified to better understand the nature of change within individual teachers' instructional practices.

Qualitative Data Analysis

To analyze qualitative classroom observation data, the original Excel database was uploaded to NVivo (version 11) data analysis software. In the first year of the evaluation, the Evaluation Team developed 9 qualitative codes deductively and 9 qualitative codes inductively. These 18 codes were applied to 2017-2018 classroom observation data, and codes were compared by lesson type for the full sample. Data from the 10 teachers with observations of all three lesson types were analyzed in depth to understand the relationship between instructional practices and lesson type. The qualitative code book can be found in Appendix C.

Findings

Descriptive Statistics of Lessons by Type

As shown in Table 2, observers observed 22 CEEMS Introductory, 33 CEEMS Challenge lessons, 14 non-CEEMS, and 28 Comparison lessons. The number of lessons observed by lesson type varied for each teacher, ranging from a minimum of 1 and a maximum of 13 observations per teacher. Observations included 31 mathematics lessons and 65 science lessons in 14 schools, across 5 districts, in Grades 6-12 classrooms. Delineated by grade band, the observations were from 55 middle school (Grades 6-8) and 35 high school (Grades 9-12) lessons. Analysis by lesson plan included 7 lessons without grade band information and 1 lesson without subject information.

Table 2. *Number of Lessons Observed by Lesson Type, Grade Band, and Subject, Inside the Classroom Observation and Analytic Protocol, 2017-2018*

Lesson Type	Grade Band	Subject			Total
		Missing	Mathematics	Science	
CEEMS Introductory	Missing	1	0	1	2
	Middle School	0	3	9	12
	High School	0	4	4	8
	Total	**1**	**7**	**14**	**22**
CEEMS Challenge	Missing	0	1	0	1
	Middle School	0	4	18	22
	High School	0	6	4	10
	Total	**0**	**11**	**22**	**33**
Non-CEEMS	Missing	0	0	1	1
	Middle School	0	2	6	8
	High School	0	1	4	5
	Total	**0**	**3**	**11**	**14**
Comparison	Missing	0	2	1	3
	Middle School	0	2	11	13
	High School	0	6	6	12
	Total	**0**	**10**	**18**	**28**
Total	Missing	1	3	3	7
	Middle School	0	11	44	55
	High School	0	17	18	35
	Total	**1**	**31**	**65**	**97**

As shown in Table 3, the average number of students in each lesson was similar across lesson types, ranging from 18 to 21 students. In addition, the classroom sizes of middle school and high school lessons were similar.

Table 3. *Classroom Size by Lesson Type and by Grade Band, Inside the Classroom Observation and Analytic Protocol, 2017-2018*

Classroom Size	*n*	*M*	*SD*	Minimum	Median	Maximum
Lesson Type						
CEEMS Introductory	21	21	5.35	9	22	29
CEEMS Challenge	31	20	5.05	8	20	28
Non-CEEMS	15	19	4.51	9	22	25
Comparison	23	18	7.35	8	15	34
Total across lesson type	90	19	5.74	8	21	34
Grade Band						
Middle School (Grades 6-8)	51	19	5.43	8	20	27
High School (Grades 9-12)	34	20	5.86	8	22	34
Total across grade band	85	19	5.63	8	21	34

Note. Due to missing information, the total numbers of observations across lesson type and across grade band are not equal.

Quantitative Comparisons by Lesson Type

The *Inside the Classroom Observation and Analytic Protocol* included individual rating-scale items and synthesis rating items for four subscales: lesson design, lesson implementation, lesson content, and classroom culture. Each subscale contained 6 to 9 individual items, with rating-scale response options ranging from 1 (“not at all”) to 5 (“to a great extent”). Response options also included “don’t know” and “Not Applicable” (response options 6 and 7), which were excluded when calculating means and standard deviations for the individual item and for subscale scores. Each subscale also contained one synthesis rating item, with response options that ranged from 1 (“not at all reflective of best practice in mathematics/science education”) to 5 (“extremely reflective of best practice in mathematics/science education”).

In addition, the classroom observation protocol included a subscale with 6 rating-scale items measuring lesson’s likely impact on students’ learning of mathematics/science. This subscale included response options that ranged from 1 (“negative effect”) to 5 (“positive effect”). Response options also included “don’t know” and “Not Applicable” (response options 6 and 7), which were excluded when calculating means and standard deviations for individual items and for subscale scores. This subscale did not have a synthesis rating item; therefore, subscale scores were calculated as mean scores of the 6 individual items.

Table 4, available in Appendix D, provides comparisons of lesson ratings and the synthesis rating, by lesson type, for the Lesson Design subscale. In general across all lesson types, most lessons were rated highly on all individual items and overall. CEEMS Challenge lessons had significantly higher ratings on incorporating tasks, roles, and interactions consistent with investigative mathematics/science; and on encouraging a collaborative approach to learning among the students than did the Comparison lessons. CEEMS Introductory and CEEMS Challenge lessons had significantly lower ratings on contributing sufficient resources to accomplishing the purposes of the instruction than did Comparison lessons. In addition, CEEMS Challenge lessons and Non-CEEMS lessons had lower ratings on providing adequate time and structure for wrap-up than did Comparison lessons. No significant differences were found for synthesis ratings for lesson design across lesson types.

Table 5, available in Appendix D, provides comparisons of lesson ratings and the synthesis rating, by lesson type, for the Lesson Implementation subscale. Ratings for the implementation of lessons were, for the most part, consistently high across lesson types. Compared to Comparison lessons, Non-CEEMS lessons had significantly lower ratings on having appropriate lesson pace for the developmental levels/needs of the students and the purposes of the lesson. Compared to Comparison lessons, both CEEMS Challenge and Non-CEEMS lessons had significantly lower ratings on using questioning strategies that were likely to enhance the development of student conceptual understanding/problem solving. No significant differences were found for synthesis ratings for lesson implementation across lesson types.

Table 6, available in Appendix D, provides comparisons of lesson ratings and the synthesis rating, by lesson type, for the Mathematics/Science Content subscale. Ratings for the mathematics/science content of lessons were, for the most part, consistently high across lesson types. Compared to Comparison lessons, CEEMS Introductory lessons had significantly lower ratings on having appropriate mathematics/science content for the developmental levels of the students in this class. Compared to Comparison lessons, CEEMS Challenge and Non-CEEMS lessons had significantly lower ratings on providing accurate content information. In addition, compared to Comparison lessons, CEEMS Introductory, CEEMS Challenge, and Non-CEEMS lessons had significantly lower ratings on engaging students intellectually with important ideas relevant to the focus of the lesson. Compared to Non-CEEMS lessons, Comparison lessons also had significantly higher ratings on having appropriate degrees of “sense-making” of mathematics/science content within this lesson for the developmental levels/needs of the students and the purposes of the lessons. No significant differences were found for synthesis ratings for lesson content across lesson types.

Table 7, available in Appendix D, provides comparisons of lesson ratings and the synthesis rating, by lesson type, for the Classroom Culture subscale. Ratings for the classroom culture of lessons were, for the most part, consistently high across lesson types. Compared to Comparison lessons, CEEMS Introductory lessons had significantly lower ratings on having a climate of respect for students’ ideas, questions, and contributions; and on having interactions that reflected collegial working relationships among students. Compared to Comparison lessons, CEEMS Introductory and Non-CEEMS lessons had significantly lower ratings on having interactions that reflected collaborative working relationships between teacher and students.

In addition, compared to Non-CEEMS lessons, Comparison lessons had significantly higher synthesis ratings for having classroom culture that reflects best practices.

Table 8, available in Appendix D, provides comparisons, by lesson type, for the Likely Impact subscale. Ratings for the likely impact of lessons on student learning were, for the most part, consistently high across lesson types. A statistically significant difference was found in students' ability to apply or generalize skills and concepts to other areas of mathematics/science, other disciplines, and/or real-life situations across lesson types, although no significant differences were found between any particular pair of lesson types. No significant differences were found for subscale scores for the Likely Impact subscale across lesson types.

Figure 1 shows the average synthesis ratings for lesson design, lesson implementation, content, classroom culture, and subscale scores for likely impact by lesson type. In general, CEEMS Introductory and CEEMS Challenge lessons had higher ratings on all aspects than did Non-CEEMS lessons; while Comparison lessons had higher synthesis ratings than did the other three types on content, classroom culture, and likely impact.

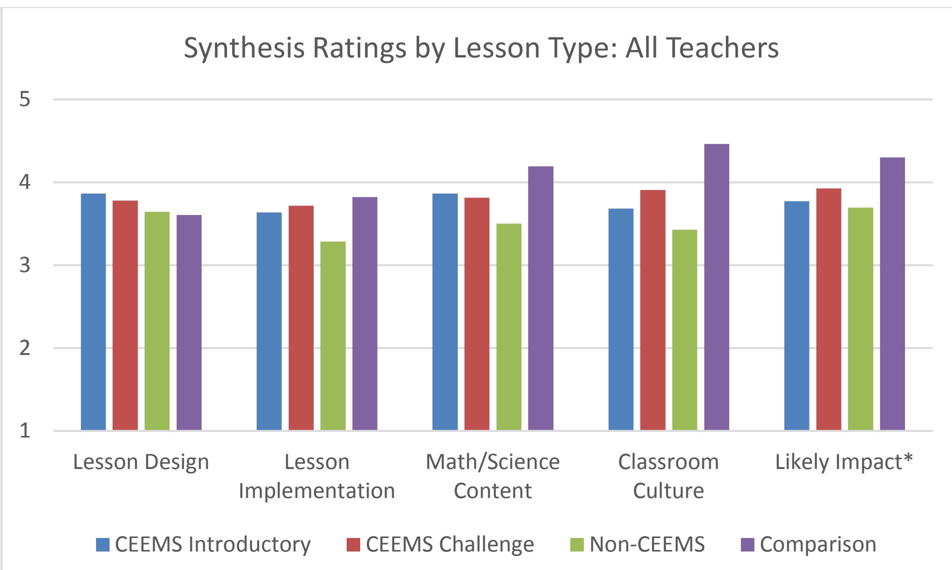

Figure 1. Synthesis rating scores by lesson type. (*Mean subscale scores were used for Likely Impact)

The classroom observation protocol allowed observers to indicate whether specific features were observed during the lesson. Table 9, available in Appendix D, provides the frequencies and percentage ratings for each lesson feature and the results of a Pearson's Chi-square test that was conducted to examine the extent to which specific lesson features were associated with different lesson types. A range of lesson features were observed in all lesson types (e.g., students completing lab notes/journals/worksheets or answering textbook questions/exercises); while other features were observed more often in certain types of lessons. Using high quality "traditional" instruction was more frequently observed in

Comparison lessons (40%) than in other types of lessons (25% in CEEMS Introductory, 22% in CEEMS Challenge, and 0% in Non-CEEMS). Teacher/students using scientific equipment was more frequently observed in CEEMS Introductory (25%), in CEEMS Challenge lessons (38%), and in Non-CEEMS lessons (29%) than in Comparison lessons (0%). In addition, CEEMS Introductory and CEEMS Challenge lessons had more than incidental reference/connection to other disciplines than did Non-CEEMS and Comparison lessons. Although not statistically significant, CEEMS Introductory, CEEMS Challenge, and Non-CEEMS lessons had more high quality "reform" instruction (e.g., investigation) than did Comparison lessons.

Quantitative Comparisons by Lesson Type for Middle School Teachers

Analyses by lesson type were conducted for middle school and high school teachers, separately. Table 10, available in Appendix D, provides comparisons of lesson ratings and the synthesis rating, by lesson type, for the Lesson Design subscale for middle school teachers. In general across all lesson types, most lessons were rated highly on all individual items and overall. Comparison lessons had significantly higher ratings on having a lesson design that reflected careful planning and organization, using instructional strategies and activities that reflected attention to students' experience, preparedness, prior knowledge, and/or learning styles, and providing adequate time and structure for wrap-up than did the other three lesson types. CEEMS Introductory lessons had significantly lower ratings on contributing sufficient resources to accomplishing the purposes of the instruction than did the Comparison lessons. In addition, CEEMS Challenge lessons and Non-CEEMS lessons had lower ratings on providing adequate time and structure for wrap-up than did the Comparison lessons. Also, compared to Comparison lessons, Non-CEEMS lessons had lower ratings on providing adequate time and structure for "sense-making" to middle school students. No significant differences were found for synthesis ratings for lesson design across middle school lesson types.

Table 11, available in Appendix D, provides comparisons of lesson ratings and the synthesis rating, by lesson type, for the Lesson Implementation subscale for middle school teachers. Ratings for the implementation of lessons were, for the most part, consistently high across lesson types. Compared to the other three lesson types, Comparison lessons had significantly higher ratings on having appropriate lesson pace for the developmental levels/needs of the students and the purposes of the lesson. CEEMS Introductory lessons had significantly lower ratings on having classroom management style/strategies that enhanced the quality of the lesson than did Comparison lessons. In addition, Non-CEEMS lessons had significantly lower ratings on teachers' confidence in their abilities to teach mathematics/science, teachers' ability to "read" the students' level of understanding and adjust instruction accordingly, and teachers' questioning strategies that enhanced the development of student conceptual understanding/problem solving than did Comparison lessons.

Table 12, available in Appendix D, provides comparisons of lesson ratings and the synthesis rating, by lesson type, for the Mathematics/Science Content subscale for middle school teachers. Ratings for the mathematics/science content of lessons were, for the most part, consistently high across lesson types. Compared to the other three lesson types, Comparison lessons had significantly higher ratings on intellectually engaging students with important ideas relevant to the focus of the lesson, on portraying mathematics/science as a dynamic body of knowledge continually enriched by conjecture, investigation analysis, and/or proof/justification, and on making appropriate connections to other areas of mathematics/science, to other disciplines, and/or to real-world contexts. Compared to Comparison lessons, CEEMS Introductory and Non-CEEMS lessons had significantly lower ratings on including elements of mathematical/science abstraction when it was important to do so and on having appropriate degrees of “sense-making” of mathematics/science content within the lessons for the developmental levels/needs of the students and the purposes of the lessons. CEEMS Introductory lessons also had lower ratings on providing appropriate mathematics/science content for the developmental levels of students than did Comparison lessons. CEEMS Challenge and Non-CEEMS lessons had significantly lower ratings on providing accurate content information than did Comparison lessons. Overall, Non-CEEMS lessons had significantly lower synthesis ratings for lesson content than did Comparison lessons.

Table 13, available in Appendix D, provides comparisons of lesson ratings and the synthesis rating, by lesson type, for the Classroom Culture subscale for middle school teachers. Ratings for the classroom culture of lessons were, for the most part, consistently high across lesson types. Compared to Comparison lessons, all three other lesson types had significantly lower ratings on encouraging and valuing active participation, on encouraging students to generate ideas, questions, conjectures, and/or propositions, and on having intellectual rigor, constructive criticism, and the challenging of ideas. CEEMS Introductory lessons had significantly lower ratings on having a respectful climate for students’ ideas, questions, and contributions than did Comparison lessons. In addition, CEEMS Introductory and Non-CEEMS lessons had lower ratings on having interactions that reflected collegial working relationships among students, as well as lower synthesis ratings for Classroom Culture than did Comparison lessons.

Table 14, available in Appendix D, provides comparisons, by lesson type, for the Likely Impact subscale for middle school teachers. Ratings for the likely impact of lessons on student learning were, for the most part, consistently high across lesson types. Compared to the other three lesson types, Comparison lessons had significantly higher impacts on students’ ability to apply or generalize skills and concepts to other areas of mathematics/science, other disciplines, and/or real-life situations. CEEMS Introductory lessons had significantly lower ratings on impacting students’ understanding of important mathematics/science concepts than did Comparison lessons. CEEMS Introductory and Non-CEEMS lessons had significantly lower ratings on impacting students’ capacity to carry out their own inquiries, on impacting students’ self-confidence in doing mathematics/science, on impacting students’ interest in and/or appreciation for the discipline, and the overall subscale scores for the Likely Impact subscale than did Comparison lessons.

Figure 2 shows the average synthesis ratings for lesson design, lesson implementation, content, classroom culture, and subscale scores for likely impact by middle school lesson type. In general, CEEMS Introductory and CEEMS Challenge lessons had higher ratings on all aspects than did Non-CEEMS lessons; although Comparison lessons had higher synthesis ratings than the other three lesson types on all aspects.

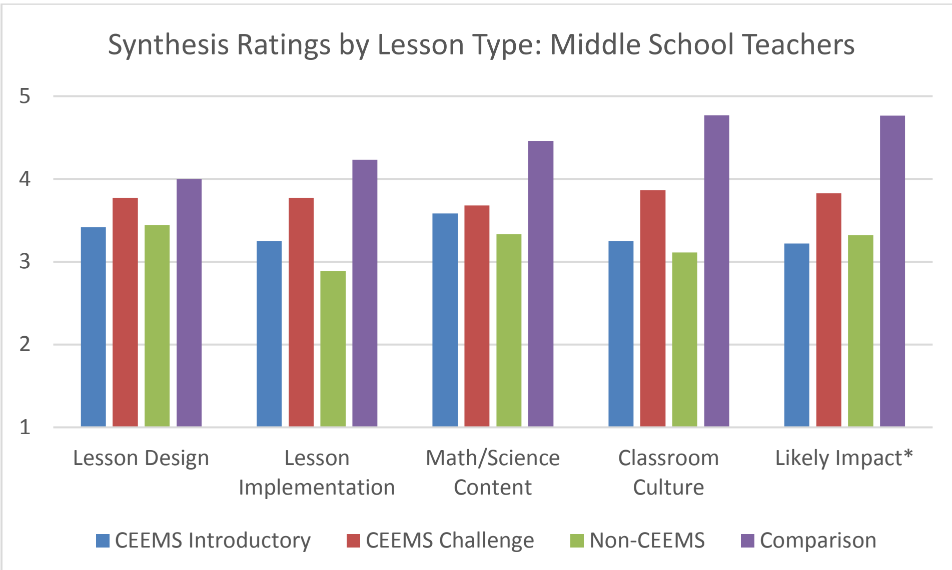

Figure 2. Synthesis rating scores by lesson type for middle school teachers. (*Mean subscale scores were used for Likely Impact.)

The classroom observation protocol allowed observers to indicate whether specific features were observed during lessons. Table 15, available in Appendix D, provides the frequencies and percentage ratings for each lesson feature and the results of a Pearson's Chi-square test that was conducted to examine the extent to which specific lesson features were associated with different lesson types. A range of lesson features were observed in all lesson types (e.g., students completing lab notes/journals/worksheets or answering textbook questions/exercises); while other features were observed more often in certain types of lessons. Teacher/students using audio-visual resources was more frequently observed in CEEMS Introductory lessons (50%) and Comparison lessons (69%) than in CEEMS Challenge (23%) and in Non-CEEMS lessons (33%).

Quantitative Comparisons by Lesson Type for High School Teachers

Table 16, available in Appendix D, provides comparisons of lesson ratings and the synthesis rating, by lesson type, for the Lesson Design subscale for high school teachers. In general across all lesson types, most high school lessons were rated highly on all individual items and overall. CEEMS Introductory lessons had significantly higher synthesis ratings for lesson design than did Comparison lessons. CEEMS Introductory, CEEMS Challenge, and Non-CEEMS lessons had significantly higher ratings on incorporating tasks, roles, and interactions consistent with investigative mathematics/science than did Comparison lessons. CEEMS Introductory and CEEMS Challenge lessons also had significantly higher ratings on encouraging a collaborative approach to learning among students than did Comparison lessons.

Table 17, available in Appendix D, provides comparisons of lesson ratings and the synthesis rating, by lesson type, for the Lesson Implementation subscale for high school teachers. Ratings for the implementation of lessons were, for the most part, consistently high across high school lesson types. Compared to Comparison lessons, CEEMS Introductory and CEEM Challenge lessons had significantly higher ratings on using instructional strategies that were consistent with investigative mathematics/science than did Comparison lessons. No significant differences were found for synthesis ratings for lesson implementation across lesson types.

Tables 18, 19, and 20 (Appendix D) provide comparisons of lesson ratings and the synthesis rating, by lesson type, for the Mathematics/Science Content subscale (Table 18), Classroom Culture subscale (Table 19), and Likely Impact subscale (Table 20) for high school teachers. Ratings for the mathematics/science content of lessons were, for the most part, consistently high across high school lesson types. No significant differences were found for individual items or synthesis ratings for lesson content across high school lesson types. Ratings for the classroom culture of lessons were, for the most part, consistently high across high school lesson types. No significant differences were found for individual items or synthesis ratings for lesson content across high school lesson types. Ratings for the likely impact of lessons on student learning were, for the most part, consistently high across high school lesson types. No significant differences were found for item ratings or for subscale scores for the Likely Impact subscale across lesson types.

Figure 3 shows the average synthesis ratings for lesson design, lesson implementation, content, classroom culture, and subscale scores for likely impact by high school lesson type. In general, CEEMS Introductory and CEEMS Challenge lessons had higher ratings on all aspects than did Comparison lessons for high school teachers.

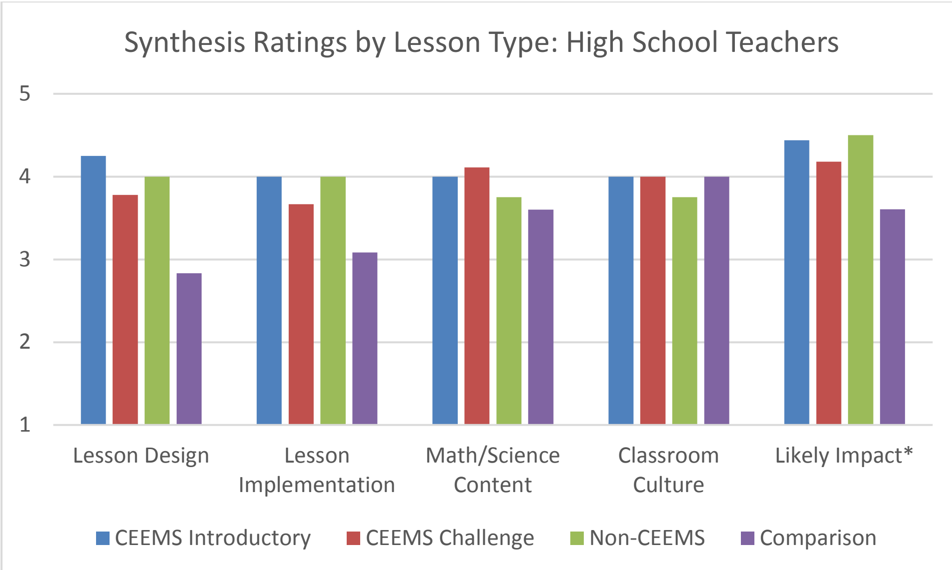

Figure 3. Synthesis rating scores by lesson type for high school teachers. (*Mean subscale scores were used for Likely Impact.)

Table 21, available in Appendix D, provides the frequencies and percentage ratings for each lesson feature and the results of a Pearson's Chi-square test that was conducted to examine the extent to which specific lesson features were associated with different high school lesson types. High quality "reform" instruction was more frequently used by CEEMS Challenge lessons (50%) in high school, compared to other types of lessons; while only Comparison lessons featured reviews/practices to prepare students for an externally mandated test (33%).

Quantitative Comparisons by Lesson Type for Matched Teacher Sample

As was discussed previously, observations of at least one of three different lesson types were provided for 10 teachers. These observation data were matched for each teacher to get a better understanding of the nature of change for individual teachers' instructional practice. Table 22, available in Appendix D, and Figure 4 provide findings for the synthesis rating scores for the matched teacher sample. All lesson types had similar synthesis ratings on lesson designs. Comparison lessons had significantly higher ratings for Mathematics/Science Content and for Classroom Culture than did Non-CEEMS lessons. Although not statistically significant, CEEMS Introductory and CEEMS Challenge lessons had higher synthesis ratings on Lesson Implementation, Mathematics/Science Content, and Classroom Culture, as well as higher subscale scores on Likely Impact than did Non-CEEMS lessons.

Figure 4. Matched teachers' synthesis scores by lesson type.

Summary of Quantitative Findings

In general, CEEMS Introductory and CEEMS Challenge lessons had higher synthesis ratings for lesson design, lesson implementation, content, and classroom culture, as well as higher subscale scores for likely impact than did Non-CEEMS lessons. Comparison lessons had higher synthesis ratings than did the other three types on content, classroom culture, and likely impact. Using high quality "traditional" instruction was more frequently observed in Comparison lessons, while teacher/students using scientific equipment was more frequently observed in CEEMS Introductory, CEEMS Challenge lessons, and in Non-CEEMS lessons than in Comparison lessons. CEEMS Introductory and CEEMS Challenge lessons had more than incidental reference/connection to other disciplines than did Non-CEEMS and Comparison lessons.

For middle school lessons, CEEMS Introductory and CEEMS Challenge lessons had higher synthesis ratings for lesson design, lesson implementation, content, and classroom culture, and also had higher subscale scores for likely impact than did Non-CEEMS lessons. Comparison lessons had higher synthesis ratings than did the other three lesson types on all aspects. Teacher/students using audio-visual resources was more frequently observed in middle school CEEMS Introductory lessons and Comparison lessons than in CEEMS Challenge and Non-CEEMS lessons.

For high school lessons, CEEMS Introductory and CEEMS Challenge lessons had higher synthesis ratings for lesson design, lesson implementation, content, and classroom culture, as well as higher subscale scores for likely impact than did Comparison lessons. High quality "reform" instruction was more frequently used by CEEMS Challenge lessons in high school, compared to other types of lessons. Only Comparison lessons featured reviews/practices to prepare students for an externally mandated test.

Using a matched teacher sample, all lesson types had similar synthesis ratings on lesson design. Comparison lessons had significantly higher ratings for Mathematics/Science Content and for Classroom Culture than did Non-CEEMS lessons for the matched teachers. Although not statistically significant, CEEMS Introductory and CEEM Challenge had higher synthesis ratings on Lesson Implementation, Mathematics/Science Content, and Classroom Culture, and also had higher subscale scores on Likely Impact than did Non-CEEMS lessons.

Qualitative Comparisons by Lesson Type

Lessons were compared qualitatively by lesson type to understand the nature of the lesson, overall.

Non-CEEMS Lessons

Across teachers, Non-CEEMS lessons were characterized by a hands-on exploratory activity, an activity that initiated students into the design part of the challenge, discussion of concepts or ideas related to the big idea, or vocabulary development and direct instruction regarding the scientific or mathematical concepts included in the challenge. Hands-on exploration included activities such as mini investigations and the use of computer simulations. Initiation into the design aspect of the challenge most often included research about how to design the object on which the challenge was centered. Discussions of concepts related to the big idea and vocabulary development/direct instruction most often included whole- and small-group discussion of scientific or engineering concepts relevant to the challenge. For four lessons in which the observer mentioned the larger CEEMS challenge, the observer also indicated that the teacher had not guided students toward development of the challenge but gave the students the challenge. Across these lessons, students often worked individually or in small groups on the activity while the teacher circulated among the students to check their progress.

CEEMS Introductory Lessons

CEEMS Introductory lessons were characterized by a number of overlapping activities that often occurred during the same class period. Most often, lecture, question-and-answer discussions, and worksheets intended to develop vocabulary or conceptual understanding were supplemented by a video related to the topic of the challenge intended to “hook” student interest, introduction to elements of engineering design, independent research about the topic, or a demonstration or hands-on activity. Many lessons included discussions of the challenge, although the teacher presented the challenge to the students more often than the students developed the specifics of the challenge through discussion and discovery. Many lessons included discussions of the big idea, essential questions, and/or guiding questions. In the lessons that included discussions of essential questions, students most often worked in groups to discuss questions they had related to the challenge, and then the teacher combined multiple discussions until a set of essential questions were agreed upon by the entire class.

CEEMS Challenge Lessons

CEEMS Challenge lessons were characterized by hands-on activities, engineering design, discussion of the challenge in small groups, presentation of a completed challenge to the class, or collection of data. For most of these lessons, students worked with everyday materials—such as cardboard, toothpicks, tape, and paper—to build a prototype of an object they designed in small groups. A few lessons were focused on testing a built prototype and/or re-designing the prototype based on the results of tests. Some lessons engaged students in the development of designs on paper, in which students were expected to demonstrate their understanding of a mathematical or scientific concept in the design. Whether designs were on paper or were physical objects, students were given certain criteria and constraints within which they were expected to work.

Comparison Lessons

Across teachers, Comparison lessons were characterized by direct instruction/lecture or whole-group direct question-and-answer discussion, completion of worksheets or reading materials individually or in small groups, or a limited amount of hands-on activities (completed by students) or demonstration activities (presented by the teacher). Lectures, whole-group discussions, worksheets, and reading materials were focused on defining concepts. Hands-on and demonstration activities were focused on visualizing scientific concepts, such as the structure of DNA or the ways the shape of a line changes in response to changes in the coordinate points or functions.

Qualitative Comparisons for the Matched Teacher Sample

For teachers who were observed during at least three of the four different lesson types (Non-CEEMS or Comparison, CEEMS Introductory, and CEEMS Challenge), elements of instruction evident in most lessons, regardless of type, were the use of hands-on materials and activities, the use of collaborative student groupings, generally high levels of student engagement, and a generally high level of attention paid to ensuring all groups and individual students were engaging in the activity and developing the expected conceptual understanding as a result. Teachers used hands-on activities and materials to demonstrate concepts in lessons of all CEEMS types and in comparison lessons. Students often worked in pairs or small groups to complete activities of the lesson, regardless of the purpose of the lesson. In general, students appeared engaged and participated actively in activities. Teachers often circulated among students to ensure their individual questions were answered, their misconceptions were clarified, and that they were developing an understanding of the concept discussed.

Elements of instruction evident in CEEMS Introductory and Challenge lessons more so than in Non-CEEMS or Comparison lessons were the use of worksheets and lecture in limited and focused ways; the use of an engineering design process that included collaboration, preliminary design, testing, and re-design during hands-on activities; the use of roles within groups; and a sustained focus on the real-world application of concepts. For CEEMS Introductory and Challenge lessons—and, to some extent, Non-CEEMS lessons—worksheets and lecture often were only one part of a more extensive lesson and were focused narrowly on specific concepts or vocabulary. In contrast, lectures and direct question-and-answer sessions or the completion of worksheets often were the primary purposes of Comparison lessons. CEEMS Introductory and Challenge lessons engaged students in several aspects of the engineering design process and required students to engage in sustained activities related to research about design, collaboration and group decision-making, scaling and labeling sketches of designs, testing of designs, and re-design based on the results of tests. In contrast, Comparison lessons that included hands-on activities almost exclusively focused on hands-on and building processes isolated from other decision-making and design processes. Many CEEMS Introductory and Challenge lessons assigned roles to students within groups, such as materials manager, electrical engineer, and recorder; these roles were not evident in Comparison or Non-CEEMS lessons. Finally, teachers focused the purpose of CEEMS units and individual lessons on a real-world topic, whereas they often mentioned real-world examples of concepts in limited ways in Comparison lessons.

Summary of Qualitative Findings

Teachers generally used a wider variety of instructional strategies, expected students to explore and research independently, used multiple elements of the engineering design process, focused lessons on real-world topics, engaged students in open-ended discussions, utilized group roles, and applied lecture and individualized seat-work (i.e., worksheets) in more focused ways more often in CEEMS and Non-CEEMS lessons than in Comparison lessons. Teachers used collaborative learning pairs and groups, used hands-on activities, encouraged engagement and active participation, and ensured all students were understanding concepts in all lessons.

Discussion and Recommendations

When triangulated, findings from focus group, quantitative classroom observation, and qualitative classroom observation data demonstrated that CEEMS participation facilitated the use of challenge-based instructional practices and the engineering design process in science and mathematics classes for grades 6-12. Teachers adopted challenge-based instructional practices and utilized the engineering design process in the context of CEEMS units. Evidence suggested that teachers used a limited number of instructional practices that were consistent with challenge-based instruction (e.g., collaborative grouping) in other units not connected to the CEEMS program. **Recommendation: to encourage future participation in CEEMS beyond the grant period, market CEEMS as an opportunity for teachers to develop skills for embedding engineering design into science classes, as is called for in the Next Generation Science Standards.**

Consistently high ratings across teachers regardless of lesson type suggested that CEEMS teachers provided high quality instruction to students in general. Although Comparison lessons were rated more highly than CEEMS lesson types on item and synthesis ratings, these higher ratings likely reflected the relative observability of student learning based on the instructional approach and the limitations of current observation protocols. In other words, it was much easier to observe the extent to which students learned the intended content when the lesson consisted of direct questioning by the teacher in a whole group setting than when students worked in groups on hands-on projects over multiple class periods. Although the *Inside the Classroom Observation and Analytic Protocol* (Horizon Research Inc., 2000) is an excellent resource for observing both traditional and inquiry-based lessons, it is limited in its ability to capture challenge-based learning and engineering design processes and must be used across multiple lessons to capture learning for multi-lesson projects. This evaluation was focused on teacher change, rather than student learning, and therefore was limited in its ability to include additional measures to account for student learning. **Recommendation: future evaluation or research on student learning about engineering design in challenge-based classrooms should include mechanisms for the observer to interact with students and/or observe enough lessons in a unit to gauge student learning.**

Higher ratings for Comparison lessons than CEEMS lesson types also may have reflected teacher learning, in that the 2017-2018 academic year was only the second full year of CEEMS implementation for cohort 5 teachers. Change in instructional practice as a result of professional learning has been found to be complex and not linear (Guskey, 2002). Therefore, high ratings for Comparison lessons suggested that additional experience implementing lessons in the CEEMS format is likely to lead to CEEMS lessons that are both high quality and inclusive of practices consistent with Next Generation Science Standards.

Conclusions

The value of the CEEMS program to changes in teacher instructional practices and Resource Team support for teachers was the change from traditional instruction to practices consistent with the Next Generation Science Standards, including engineering design, collaborative grouping, critical thinking, and connection of science and mathematics content to real-world issues. This section responds to the evaluation questions based on findings from all three years of the external evaluation.

In what ways did teachers' instructional practices change in the course of their participation in CEEMS?

Teachers used a wider variety of instructional practices in CEEMS lessons than in traditional lessons. Instructional practices evident in CEEMS lessons included probative, open-ended questioning that encouraged critical thinking; the engineering design process; challenge-based learning strategies; collaborative grouping; and external resources (e.g., videos) as a means to focus the lesson on real-world issues.

Teachers supported their instructional practices by drawing upon a network of other CEEMS teachers and Resource Team members that emerged as a direct result of the CEEMS program. This network generated a vast amount of curriculum resources, such as lesson plans, that were shared and made available to the wider public electronically.

Change in instructional practice was dependent upon individual effort and inclination toward change and was not guaranteed to result from participation.

In what ways did CEEMS Resource Team support for teachers change in the course of their participation in CEEMS?

The Resource Team developed and improved upon a set of resources—such as instructional guides—and processes—such as post-unit debriefing—through trial and error over the years of the program. These improvements enhanced the teachers' experiences by facilitating and streamlining communication and providing a wealth of resources that had been applied in real classroom settings. The CEEMS project team's willingness to incorporate suggestions and changes initiated by the Resource Team enabled these improvements.

The Resource Team developed a collaborative coaching structure in which multiple Team members had multiple contacts with individual teachers over time. The collaborative structure allowed Resource Team members to learn from one another and draw upon a common set of resources.

The Resource Team developed a number of coaching strategies to support instruction that was responsive to a challenge-based learning environment. Some of the most important coaching strategies supported teachers' ability to develop new challenge-based engineering design units and lessons consistent with grade-level standards, ask probing questions that sparked critical thinking, ensure science/mathematics content learning occurred alongside engineering design activities, manage a classroom environment in which students worked in small groups on hands-on activities, and use student work to connect with building administrators in positive ways.

Individual Resource Team members who were former engineers discussed engineering as a career within classrooms, thus supporting student understanding of engineering in the real world.

References

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching: theory and practice*, *8*(3), 381-391.

Horizon Research Inc. (2000). *Inside the classroom observation and analytic protocol*. Retrieved from http://www.horizon-research.com/horizonresearchwp/wp-content/uploads/2013/04/cop1.pdf.

Maltbie, C. & Butcher, K. (2014). *NSF MSP: CEEMS project – Year 3 complete academic year: Evaluation summary report.* Cincinnati, OH: Evaluation Services Center.

Woodruff, S. B., Dixon, M., & Li, Y. (2016). *Evaluation of CEEMS: The Cincinnati Engineering Enhanced Mathematics and Science: Annual report 2015-2016.* Oxford, OH: Miami University, Discovery Center for Evaluation, Research, and Professional Learning.

Woodruff, S. B., Dixon, M. L., & Li, Y. (2017). *Evaluation of CEEMS: The Cincinnati Engineering Enhanced Mathematics and Science Partnership Project: Annual report, 2016-2017.* Oxford, OH: Miami University, Discovery Center for Evaluation, Research, and Professional Learning.

Appendices