

## **AC 2010-1680: SPECIAL SESSION: MODEL-ELICITING ACTIVITIES: A CONSTRUCT FOR BETTER UNDERSTANDING STUDENT KNOWLEDGE AND SKILLS**

### **Tamara Moore, University of Minnesota**

Tamara J. Moore is an Assistant Professor of Mathematics/Engineering Education and co-director of the STEM Education Center at the University of Minnesota. Dr. Moore is a former high school mathematics teacher and her research interests are centered on the integration of STEM concepts through contextual problem solving in the mathematics and engineering classroom. She has been developing curricular tools and researching professional development and student learning in this area. Before coming to the University of Minnesota, Dr. Moore received her Ph.D. from the School of Engineering Education at Purdue University.

### **Brian Self, California Polytechnic State University**

Brian Self is a Professor in the Mechanical Engineering Department at California Polytechnic State University in San Luis Obispo. Prior to joining the faculty at Cal Poly in 2006, he taught for seven years at the United States Air Force Academy and worked for four years in the Air Force Research Laboratories. Research interests include active learning and engineering education, spatial disorientation, rehabilitation engineering, sports biomechanics, and aerospace physiology. He worked on a team that developed the Dynamics Concept Inventory and is currently collaborating on a grant to develop and assess Model Eliciting Activities in engineering. Brian is the 2008-2010 ASEE Zone IV Chair and serves as Cal Poly's ASEE Campus Representative.

### **Ronald Miller, Colorado School of Mines**

Ronald L. Miller is professor of chemical engineering and Director of the Center for Engineering Education at the Colorado School of Mines where he has taught chemical engineering and interdisciplinary courses and conducted engineering education research for the past 24 years. Dr. Miller has received three university-wide teaching awards and has held a Jenni teaching fellowship at CSM. He has received grant awards for education research from the National Science Foundation, the U.S. Department of Education FIPSE program, the National Endowment for the Humanities, and the Colorado Commission on Higher Education and has published widely in the engineering education literature.

### **Margret Hjalmarson, George Mason University**

Margret A. Hjalmarson is an Assistant Professor in the Graduate School of Education at George Mason University. She received her B.A. in mathematics at Mount Holyoke College, and her Masters mathematics and Ph. D. in mathematics education at Purdue University. She currently coordinates the Mathematics Education Leadership program for K-8 mathematics specialists and Ph.D students in mathematics education. She works with engineers to study the learning of mathematics in engineering education contexts. She also works in mathematics teacher professional development and design research related to students' learning of mathematics.

### **Judith Zawojewski, Illinois Institute of Technology**

Judith S. Zawojewski is an Associate Professor in the Department of Mathematics and Science Education at Illinois Institute of Technology in Chicago, IL. She received her B.S. in mathematics and education at Northwestern University, and her Masters and Ph. D. degrees in mathematics education at National-Louis University and Northwestern University respectively. She works with teachers of mathematics and engineering educators who are interested in improving the learning and teaching of mathematics in their educational environment. In particular, she is studies the role of modeling and problem solving in developing mathematical capabilities.

**Barbara Olds, Colorado School of Mines**

Barbara M. Olds is Associate Provost for Educational Innovation and Professor of Liberal Arts and International Studies at the Colorado School of Mines. She returned to CSM in 2006 after spending three years at the U. S. National Science Foundation where she served as the Division Director for the Division of Research, Evaluation and Communication (REC) in the Education and Human Resources Directorate. She remains a consultant to the EHR Directorate. During the 2006-2007 academic year Barbara was a visiting professor in Purdue University's Engineering Education Department. Her research interests are primarily in understanding and assessing engineering student learning. She has participated in a number of curriculum innovation projects and has been active in the engineering education research and assessment communities. She is a Fellow of the American Society for Engineering Education and was a Fulbright lecturer/researcher in Sweden.

**Heidi Diefes-Dux, Purdue University**

Heidi A. Diefes-Dux is an Associate Professor in the School of Engineering Education at Purdue University. She received her B.S. and M.S. in Food Science from Cornell University and her Ph.D. in Food Process Engineering from the Department of Agricultural and Biological Engineering at Purdue University. Since 1999, she has been a faculty member within the First-Year Engineering Program at Purdue, the gateway for all first-year students entering the College of Engineering. She coordinated (2000-2006, 2010) and continues to teach in the required first-year engineering problem solving and computer tools course, which engages students in open-ended problem solving and design. Her research focuses on the development, implementation, and assessment of model-eliciting activities with realistic engineering contexts. She is currently the Director of Teacher Professional Development for the Institute for P-12 Engineering Research and Learning (INSPIRE).

**Richard Lesh, Indiana University**

Richard Lesh, Professor of Learning Science and Director of Center for Research on Learning and Technology at Indiana University, is the creator of Model-Eliciting Activities and a leading researcher in the area of implementing models and modeling in STEM curriculum – particularly in mathematics, research methodologies, and complex assessment systems.

# Model-Eliciting Activities: A Construct For Better Understanding Student Knowledge and Skills

## Introduction

The ABET criteria for engineering programs include that students should have “an ability to apply mathematics, science and engineering”, “an ability to design a system, component, or process to meet desired needs”, “an ability to identify, formulate and solve engineering problems”, and “an ability to communicate effectively”, and “a knowledge of contemporary issues”<sup>1</sup>. One manner of integrating teamwork and engineering contexts in undergraduate engineering is through the educational construct of Model-Eliciting Activities (MEAs). MEAs are a class of interdisciplinary problems designed to simulate authentic, client-driven situations in classroom settings. MEAs allow teachers and researchers to observe student development of conceptual models by requiring students to make their models explicit through design-test-revise cycles. The solution of an MEA requires the use of one or more mathematical or engineering concepts that are unspecified by the problem - students must make new sense of their existing knowledge and understandings to formulate a generalizable mathematical model that can be used by the client to solve the given and similar problems. An MEA creates an environment in which skills beyond mathematical abilities are valued because the focus is not on the use of prescribed equations and algorithms but on the use of a broader spectrum of skills required for effective engineering problem solving. Carefully constructed MEAs can begin to prepare students to communicate and work effectively in teams; to adopt and adapt conceptual tools; to construct, describe, and explain complex systems; and to cope with complex systems. MEAs provide a learning environment that is tailored to a more diverse population than typical engineering course experiences as they allow students with different backgrounds and values to emerge as talented, and that adapting these types of activities to engineering courses has the potential to go beyond “filling the gaps” to “opening doors” to women and underrepresented populations in engineering. Further, MEAs provide evidence of student development in regards to ABET standards. MEAs are particularly useful for implementation in engineering training as they promote creative problem solving, application of interdisciplinary knowledge, and teamwork.

This paper will present four cases of research on student learning through MEAs developed and assessed through an NSF-funded grant, *Collaborative Research: Improving Engineering Students’ Learning Strategies Through Models and Modeling*. We have added a secondary title, *Modeling: Elicitation, Development, Integration, and Assessment (MEDIA) Project*, to more easily describe the work that we are doing. The MEDIA Project is a large-scale, four-year collaborative research project between seven major universities: University of Pittsburgh, University of Minnesota, US Air Force Academy, Colorado School of Mines, Purdue University, Pepperdine University, and California Polytechnic State University, through which we are working to develop, assess, and evaluate MEAs in undergraduate engineering courses, especially focusing on second and third year courses.

## Description and Assessment of MEAs

Model-Eliciting Activities (MEAs) are problems in which a client requests that an engineering system (e.g., a procedure for maximizing output, a mathematical model, a process to analyze data, a ranking method) be created, explained, or adapted to meet specific needs of a situation. When engineering education problems are explicitly designed to elicit a model from the student(s), little time is wasted in students pursuing irrelevant side trips often associated with open-ended problems, yet there is enough that a variety of reasonable systems can be devised, explained, or adapted to meet the given constraints. The MEA brings an aspect of the ABET<sup>1</sup> criteria alive: “design-to-constraints.”

One of the main reasons for using MEAs in engineering education is that they are designed to be thought revealing. Because the answer to the problem is the creation, explanation or adaptation of an engineering system, students’ thinking is articulated publicly. This thought-revealing characteristic provides opportunities for student self-assessment and reflection, as well as peer’s public evaluation of each other’s thinking. Further, teachers are in a position to gain authentic knowledge of how students are thinking—which informs subsequent instruction. In particular, the nature of students’ conceptions of foundational engineering constructs is readily accessible to the teacher, as well as the researcher.

Assessment of students’ responses to MEAs can take on two forms. One means for assessing student work is to describe the characteristics and nature of the models students create in response to an MEA. Carmona<sup>2</sup> produced a system for describing responses to MEAs, and Hjalmarson<sup>3</sup> has adapted the system to describe work in engineering-based MEAs. The result provides information that reveals how students are thinking about and using engineering and modelling ideas. The other type of assessment is evaluative, wherein one attempts to put a value on how well the system created, explained, or adapted in order to meets the needs of the client. Diefes-Dux, Zawojewski and Hjalmarson<sup>4</sup> describe the difficulties associated with evaluative assessment. A major challenge is to identify criteria for evaluating students’ solutions that reflect what would be valued in the professional engineering environment. Designing valid and reliable tools for use by multiple instructors to evaluate work in consistent ways is another challenge, especially when a variety of reasonable solutions may be produced<sup>5</sup>. It is also important and challenging to devise a system that can motivate student learning by providing formative feedback to students<sup>6</sup>.

Diefes-Dux, et al.<sup>4</sup> present a concrete example of how educational design research, a models-and-modeling perspective from mathematics education, and multi-tiered teaching experiments have been used in the design of valid and reliable evaluation tools for scoring team responses to MEAs. Their work demonstrates how the design of a package of evaluation tools (including rubrics, task-specific supports, and scorer training) based on the aforementioned educational research methods supports (1) sustained fidelity to engineering expert-identified characteristics of high performance across iterations of change to improve reliability, and (2) the implementation of planned iterations of the evaluation tools based on systematically collected data. Embedded in a system where students, teaching assistants and instructors use a common rubric to generate iterations of peer assessment, TA feedback, and final evaluation for a grade. The rubric for every MEA addresses: (1) the appropriateness of the model generated, explained

or modified, (2) the generalizability of the model so it can be used in similar situations or readily adapted to slightly different situations, and the (3) share-ability of the model so that it communicates readily with the intended client. By developing clear definitions of each of these characteristics of a good response, and developing problem-specific guidelines for the final evaluation phase, Diefes-Dux, et al.<sup>4</sup> demonstrate how using design study methodology to develop the assessment tools supports (1) sustained fidelity to engineering expert-identified characteristics of high performance across iterations of change to improve reliability, and (2) the implementation of planned iterations of the evaluation tools based on systematically collected data.

## **Method: Case Study**

A multiple-case evaluative case study method<sup>7</sup> was used to look at student learning during MEAs for this paper<sup>7, 8</sup>. According to Creswell<sup>9</sup>, the purpose of a case study is to explore a program (or other entity) in depth. “The cases are bounded by time and activity, and researchers collect detailed information using a variety of data collection procedures over a sustained period of time.” In this paper, each case is an implementation of an MEA at a different research site within the MEDIA Project. The researcher in charge of implementation used a method for analyzing the impact of student learning that fit his/her case. These methods will be described within each case. The first two cases, Accident Reconstruction MEA and Wet Suit MEA, focus on using pre-post analysis to look at student learning, particularly using concept inventories. The second two cases, NanoRoughness MEA and NASA Advanced Life Support MEA, focus on looking at student responses to the activity in depth.

## **The Cases**

### ***Accident Reconstruction MEA***

This MEA targets the principles of particle work-energy, impulse momentum, and impact in a sophomore-level dynamics class. A major concept addressed in the MEA is that mechanical energy is lost during an impact. Simple work-energy problems involving accidents are given as homework assignments prior to the MEA, and then a short pre-read provides background on some of the basics of accident reconstruction. For the MEA, the new Traffic Division in Sri Lanka has asked the student teams to develop a set of guidelines and procedures to use at an accident site for determining what actually happened during the crash (how fast was the car going, etc). They provide two different accident cases to guide the students into creating their initial guidelines, and then send an additional two scenarios so that the students can test and refine their procedures. The accident cases were obtained from an actual police department and contain a lot of extraneous information – students must make different simplifying assumptions, determine which information is relevant, and analyze each of the accidents to determine if the drivers were speeding. Additionally, they are required to make a recommendation on if the driver should be prosecuted or not; this will bring in some decisions based on confidence in their calculations and how they think their assumptions affect the certainty of their results

The principles of particle work-energy and impulse momentum are often discussed in terms of perfect spheres and abstract shapes that have little to do with realistic engineering scenarios. In

order to provide some engineering context to these principles, we have developed an Accident Reconstruction MEA. In this MEA, student teams are asked to create a guidebook and procedures for the new Traffic Police Division in Sri Lanka. By utilizing the principle of work energy before and after the impact, as well as impulse momentum through the impact, the students should be able to determine the approximate speed of the vehicle(s) involved in the crash.

Before the full MEA was implemented, students were given several homework problems involving simple collisions. This provided them with some initial feedback on how to approach fairly easy problems before going on to more complex scenarios. For the MEA, actual accident reports from a local police station were used to generate different “cases”. Using the homework problems and two of these cases, the students developed a draft set of accident reconstruction guidelines. The Sri Lankan Senior Deputy Inspector General of Police then sent the teams two additional accident cases so that the students could determine if their procedures were sufficient. For the four accidents, the students were asked to make a judgment on if the drivers were speeding, and if they should be prosecuted. For this last task, the students had to consider the certainty of their findings and how much error might be involved in their calculations. The project seemed to be well received based on the results of a survey. See Table 1 for results.

Table 1. Summative results from student responses to a survey on the Accident Reconstruction MEA.

	Strongly agree	Agree	Neither Agree nor Disagree	Disagree	Strongly disagree
The Accident reconstruction project helped me learn the material	6	34	14	8	1
The Accident reconstruction project was interesting and motivating.	2	30	17	13	1

Additional open-ended questions were also answered by 90 students. Themes and the number of times they were mentioned are provided below:

*What did you like about the Project and why?:*

Practical application, real world (54), Group activity (22), Helped me learn the concepts (16), Had to make assumptions (6), Applied multiple concepts (3), Allowed us to be creative (1), Focused on process not answers (1).

*What didn't you like about the Project and why?:*

Vagueness of assignment and/or scenarios (25), Group difficulties (including meeting times outside of class) (15), Writing a memo (8), Simplifying the procedure to laymen's terms (6), No example answers provided (3)

Finally, we administered the Dynamics Concept Inventory to students in five sections that utilized two different MEAs (149 students) and three sections that did not (80 students); the sections were taught by different instructors. The normalized gain for the MEA sections was 29.6, and was 21.1 for the non-MEA sections. When only the two questions on the DCI that

were related to impact (and the Accident Reconstruction MEA) were examined, the MEA group had a 41.1% normalized gain, compared to 14.8% for the non-MEA group. These results indicate that the Accident Reconstruction MEA may have had a positive effect on the conceptual understanding of our dynamics students.

The MEAs were graded on four areas: (1) Was the cover memo an effective communication to the client that addresses their needs? (2) Were the procedures generalizable, usable by policemen, and easy to follow? (3) Did they apply the procedures to the last two accidents? (4) Was their approach to analyzing each of the four accident scenarios correct?

Out of the 18 total groups, only two failed to really develop a set of procedures for the MEA. These groups basically solved the four accident scenarios as if they were just four homework problems and then turned them in with a poor attempt at a cover memo. In general, the groups approached the solutions to the accident cases well. Only two teams failed to recognize that impulse and momentum are vector quantities and applied the relationships incorrectly. Four of the groups did not do a good job of developing a generalizable procedure and created separate procedures for each of the four accident cases. Seven of the groups did an excellent job of explicitly applying their procedure to the final two accident scenarios. As shown by many of the student comments and evident in their submissions, many of the teams struggled with two things: writing the procedures for a customer who is a non-engineer and creating a procedure that is generalizable to all accident scenarios.

### ***Wet Suit MEA***

The Wet Suit MEA requires student teams work for a research and design division of a wetsuit company that wants to extend their business by developing new wetsuits to market to users that might normally chose a dry suit for their needs. The teams produce a procedure for the company to estimate the time a user can stay in the water using a wetsuit made of a certain type and thickness of material. This will allow the company to have an initial performance screening for new materials without incurring cost to create a prototype model.

The Wetsuit MEA has been piloted in three upper-division engineering classes [2 transport classes at one university and 1 bioengineering transport class at a different university]. Over 125 students in groups of 3-4 have participated in solving this MEA. Wetsuit asks students to develop a transient heat transfer model for predicting the onset of mild hypothermia in divers and surfers as a function of water temperature, level of physical exertion, human body size and mass, and wet suit properties and dimensions. To successfully create this model, students must draw on their knowledge of energy balances from thermodynamics and heat transfer courses and they must make decisions about the geometry to use (usually cylinder or flat plate), about heat transfer resistances to consider and those to ignore, about how to simplify the temperature profile in the human body, and about how to estimate the rate of heat generation in a human as a function of exertion rate. Successful groups obtain an ordinary differential equation relating body temperature to time in the water that then can be analytically or numerically solved for specified initial conditions. Models that are carefully developed and solved will agree within engineering accuracy (~10-20%) for available literature data on hypothermia onset.

As part of evaluating the impact of Wetsuit, students were asked to answer 14 concept questions focused on rate vs. amount of heat transfer and steady-state vs. thermal equilibrium processes. Pre/post scores increased from 6.1 to 7.2 correct answers for bioengineering transport students, a small but statistically significant effect while scores for non-bioengineering transport students increased from 7.3 to 7.5 correct answers, a statistically insignificant amount. Although based on limited data, these results suggest that Wetsuit does not significantly impact students' conceptual understanding of heat transfer concepts – not a surprising result, since Wetsuit was not designed to focus on student misconceptions per se. In addition, pre-test raw scores were slightly higher for the non-bioengineering transport students, perhaps as a result of heat transfer study prior to solving Wetsuit (the students in the non-bioengineering transport phenomena take an earlier course in heat transfer while the bioengineering transport students have not studied heat transfer prior to taking bioengineering transport phenomena).

Anecdotally, instructors at both institutions noted that students working on Wetsuit needed more coaching to get started on this MEA than expected but once they got started, most groups did well. Students were generally motivated by the “non-textbook” nature of the MEA and several showed unexpectedly high levels of creativity in their modeling work.

### ***NanoRoughness MEA***

The NanoRoughness MEA requires that student teams develop a procedure to measure roughness given Atomic Force Microscope (AFM) images of three different samples of gold. The students have previously been given background reading on the function of the AFM. The motivation for developing the procedure is established by using a realistic context in which a company specializing in biomedical applications of nanotechnology wishes to start producing synthetic diamond coatings for joint replacements. The company intends to extend its experience with gold coatings for artery stents to this new application. The teams must create a procedure to measure roughness using images of gold because the company currently only has one image of diamond since diamond coatings are still in development. Student teams of three or four are required to establish a procedure for measuring the roughness of gold samples that could be applied to diamond samples as they are created. The students then apply the procedure to three different samples of gold. The team must write a memo to the company describing their procedure and its application to the sample AFM images and listing the additional information needed to improve their procedure.

The goal of the MEAs is for students to use the pixel values from the images as a measure of the height of the surface at any given point. Students then need to determine a procedure for determining the roughness. This leads naturally to the use of mean and standard deviation for establishing a typical height and the variation in the height. However, the first challenge for students is to find a sample of points from the surface since there is no obvious data set. One complication to this process was that the three sample images were not the same scale (ideally, to prevent the students from “eyeballing” the roughness and forcing them to take some measurements). The process of first finding a sample and then analyzing descriptive statistics from the sample is an informal inference process for determining which surface is the most rough. The NanoRoughness MEA was implemented in a large-scale (N~1400 students, or 350 teams), required first-year engineering course.



### *Student Strategies*

Responses from 35 teams were coded for the types of strategies students used to determine the roughness of the surface. Hjalmarson<sup>10</sup> and Moore & Hjalmarson<sup>11</sup> described the analysis process for sorting categories of student work. For sampling, students' procedures fall roughly into 4 categories: drawing random points or lines, using the whole image, drawing a grid on the image, and finding a random area or cross-section. Most students made an adjustment in their procedure to account for the difference in scale between the three sample images. Overall, students seemed to understand the need to find a random sample in order to compare the images. However, they did not have a rationale for a number of data points to use in their sample. Drawing random lines and then using points from the lines on fixed intervals is a common procedure used by engineers to measure the roughness of the surface of a material so some teams also came close to typical, authentic engineering procedures.

For the second part of their procedure, the descriptive statistics, the student teams typically found the mean (22 teams) and the standard deviation (23 teams) in order to differentiate the roughness of the samples. A few teams also used the median, mode, range or extreme values (maximum and minimum) in order to characterize the roughness of the surface. In some cases, tiebreaker statistics were used if the first statistic calculated did not differentiate sufficiently between the images. Use of the mean and the standard deviation was also typical in another statistics-focused MEA<sup>3</sup>. The complex part of the task is not finding the statistics (most teams used Excel® or MatLab®) but in first deciding what to calculate and they interpreting the findings in light of the context. For instance, does a rougher surface have more peaks, many peaks of similar height, or greater variation in a few peaks?

### ***NASA Advanced Life Support MEA***

The NASA Advanced Life Support MEA requires students to create a procedure for analyzing life support systems in order to determine which system is best for use in space. The teams are given a set of data for five systems that experts from NASA have already ranked from best to worst. Team must use this data to develop a mathematical procedure that uses the data provided to rank these five systems. When the teams' procedures are applied to the NASA ranked data, the systems should be ranked in the same order. The students are then asked to apply their procedure to a new set of specifications for life support systems in order to rank them from best to worst.

This MEA was implemented in a large first-year required engineering course. A task analysis was performed to assess student work on the NASA MEA. This research is described in more detail by Wang, Moore, Plumb, and Roehrig<sup>12</sup>. Here, we assessed 100 student teams using the Quality Assurance Guide (QAG), a rubric designed to assess the quality of overall team responses to MEAs. The QAG (displayed in Table 2), provides a general holistic rubric for evaluating the extent to which a team's solution meets the client's needs<sup>13</sup>. This evaluation tool has been used in many research studies involving MEAs<sup>2, 14-19</sup>.

Table 2. Quality Assurance Guide (QAG).

To prepare to assess quality of the solution (mathematical model), put yourself in the role of the client.

To do this, it's necessary to be clear about answers to the following questions:

- Who is the client?
- What solution (mathematical model) does the client need?
- What does the client need to be able to do with the solution (mathematical model)?

Then, the quality of solution can be determined by focusing on the question:

***How useful is the solution (mathematical model) for the purposes of the client?***

Quality Score	Performance Level	How useful is the solution (mathematical model)?
1	Requires redirection	The product is on the wrong track. Working longer or harder won't work. The students may require some additional feedback from the instructor.
2	Requires major extensions or revisions	The product is a good start toward meeting the client's needs, but a lot more work is needed to respond to all of the issues.
3	Requires editing and revisions	The product is on a good track to be used. It still needs modifications, additions or refinements.
4	Useful for this specific data given, but not shareable and reusable OR Almost shareable and reusable but requires minor revisions	No changes will be needed to meet the immediate needs of the client for this set of data, but not generalized OR Small changes needed to meet the generalized needs of the client.
5	Sharable or reusable	The solution not only works for the immediate situation, but it also would be easy for others to modify and use it in similar situations.

The students' product should make it clear that:

- The students went beyond producing a solution that *they* themselves can use to also produce a solution that *others* can use – by including needed explanations, and by making it as simple, clear and well-organized as possible.
- The students went beyond thinking *with* the solution to also think *about* it – by identifying underlying assumptions (so that others know when the solution might need to be modified for use in similar situations)
- The students went beyond *blind* thinking to also think *about* their thinking (by recognizing strength and weaknesses of their approach compared with other possible alternatives).

Secondly, a task model was created to represent the subtask for strategy deployment, as well as to specify shallow and deep strategies utilized by student teams in these areas. The development of this task model involved utilization of tools from the Applied Cognitive Task Analysis (ACTA) method<sup>20</sup>, a streamlined set of cognitive task analysis (CTA) tools. This was to understand what types of subtasks students needed to complete to create a quality solution to the NASA MEA. In this case, the ACTA subtasks identified for the NASA MEA were: (1)

determine initial ranking strategy for factor values; (2) apply ranking strategy to factor values; (3) determine a weighting system for each factor; (4) apply the weighting strategy for each factor; and (5) determine the final ranking of the life-support subsystems based on rankings and weightings of factors. Student work products were assessed in order to determine (a) whether or not the student team solutions addressed the five subtasks, and (b) the strategies that students employed within each subtask.

In all 100 student work products, we identified strategies for each of the 5 subtasks. Within each subtask, we identified 3-6 different specific strategies employed by student teams in their work products. Deep and shallow strategies in each of the 5 subtask areas were determined by considering aspects of expertise and cognitive difficulty.

Comparisons of deep and shallow groupings in each subtask indicate significant differences in QAG Score for 3 subtasks - *Determine a Weighting System*, *Apply Weightings*, and *Determine Final Rankings*. There was no statistically significant difference in Overall Score between groups that applied deep and shallow strategies in *Determine factor rankings* and *Apply Factor Rankings*.

Ordinal regression of QAG Score on all subtasks coded for deep and shallow strategies produced a pseudo R<sup>2</sup> of .529 (Cox-Snell). The Determine Ranking and Apply Ranking categories did not have statistically significant unique relationships to the dependent variable when removed from the model ( $p=.151$  and  $.199$  respectively). Likewise, ordinal regression of Overall Score on each individual subtask did not indicate statistically significant relationships for Determine Ranking and Apply Ranking ( $p=.063$  and  $.494$  respectively). In both of these analyses, the use of deep strategies in other three subtasks had a significant relationship with the QAG Score.

The results of the task analysis revealed that this MEA focused on two particular skills: finding patterns in data and using mathematical procedures to manipulate data. Student teams that successfully employed both of these skills created robust models for the given task. The evidence for this comes from exploring the cognitive demands of each subtask to determine why particular tasks were difficult and what cues and strategies are used to complete each subtask. In each case, the deep strategies were difficult because they involved either noticing patterns in the data given in the activity or complex manipulation of the data using mathematical procedures. Another skill indirectly involved in this activity is the ability to judge the applicability of procedures across multiple situations. This includes being able to simplify procedures in order to make them more robust. Student teams that were most successful in this MEA were able to consider how their model would apply across different circumstances.

While it is important to identify which skills and abilities are elicited during the activity, it can be just as informative to know which are not elicited during the activity. By conducting a task analysis, we discovered that no content-specific knowledge about life-support systems or space travel was needed to complete this task. This does not necessarily identify a weakness of the activity; the skills and abilities involved in the activity should be judged in relation to the larger curricular goals of the course. However, if a goal of the curriculum is to have students gain and apply knowledge of life-support systems and/or space travel to an engineering task, the activity should require students to do so.

## Discussion

The four cases of MEA development show a variety of ways to use MEAs as an assessment tool in engineering courses. The first two cases showed more instructor friendly versions of assessment, while the second two cases were more researcher oriented. Each type of assessment has merit in determining the robustness of student solutions. It is important to note, however, that MEAs were not originally designed to be instructional tools. They were initially created as research sites. As research tools, MEAs are designed to document thinking (because the models produced and the assumptions articulated that underlie the model are thought revealing), and thus provide ideal settings to assess the knowledge and abilities of teams. While full knowledge of students' thinking during the full solution process is not necessarily documented (unless observing teams, or video recording their iterations), the final solution to the problem reveals how the team finally thought about the modeled situation and how the team mapped that situation to a model they created.

Instructors who wish to use MEAs in their classrooms should use MEAs as up front investigative tools to find out what students understand and where they struggle. Because MEAs are thought revealing, instruction after an MEA can be shaped by what has been revealed in the student solutions. Here, building upon prior student knowledge is key. Using MEAs as formative assessment means that instructors will need to interpret the evidence of student understanding elicited from the MEA, use an interpretive framework that articulates the subgoals of the task (such as the Quality Assurance Guide), and use the information to provide feedback to the students regarding their understanding and feedback to the instructor on where to go next in terms of instruction. This promotes a shared understanding and ownership of the learning goals and further promotes student self-assessment.

## Conclusions

This paper presents results from the evaluation of student work on several MEAs implemented in undergraduate engineering courses. We analyzed work products from these MEAs to explore the relationship between the problem-solving strategies employed during the activity and the quality of the conceptual models, assessed by different methods for each MEA. This work provides information for instructors on the type of feedback to provide students engaging in the activity, what to do next in class, as well as validation of holistic assessments of student work. This analysis also has implications for determining the specific learning that occurs during a complex problem-solving activity, including whether or not the learning is robust.

Even though MEAs were not created as instructional tools, they have been shown to be effective as a method to allow students to learn modeling and deep engineering concepts in undergraduate engineering courses. The characteristics of MEAs, complex problem solving tasks set in realistic contexts with clients, place MEAs in the authentic assessment category<sup>21, 22</sup>. Solutions to MEAs are generalizable models, which reveal the thought processes of the students. The models created include procedures for doing things and more importantly, metaphors for seeing or interpreting things. The activities are such that student teams of three to four express their mathematical model, test it using sample data and revise their procedure to meet the needs of their client. The MEA framework provides a means to not only deliver more open-ended engineering problems

(engineering content) but also address multiple ABET criteria, especially those that are problematic to integrate in engineering courses<sup>23</sup>.

## Acknowledgment

This research is supported in part by the National Science Foundation through projects 0717801, 0717529, 0717508, 0717864, 0717751, 0717595: “Collaborative Research: Improving Engineering Students’ Learning Strategies through Models and Modeling.”

## Bibliography

1. ABET, "Criteria for accrediting programs in engineering," ABET, Inc., Baltimore, MA2004.
2. G. Carmona, "Designing an assessment tool to describe students’ mathematical knowledge," West Lafayette, IN: Purdue University, 2004.
3. M. Hjalmarson, "Engineering students designing a statistical procedure," *Journal of Mathematical Behavior*, vol. 26, pp. 178-188, 2007.
4. H. A. Diefes-Dux, J. S. Zawojewski, M. A. Hjalmarson, and "Using educational research in the design of evaluation tools for open-ended problems," *International Journal of Engineering Education*, in press.
5. D. C. Davis, G. K. L., M. S. Trevisan, and D. E. Calkins, "Engineering design assessment process and scoring scales for program improvement and accountability," *Journal for Engineering Education*, vol. 91, pp. 211-221, 2002.
6. P. Black, C. Harrison, C. Lee, B. Marshall, and D. William, *Assessment for learning: Putting it into practice*. Berkshire, England: Open University Press, 2003.
7. R. K. Yin, *Case study research: Design and methods*, 3rd ed. Thousand Oaks, CA: Sage Publications, 2003.
8. P. Darke, G. Shanks, and M. Broadbent, "Successfully completing case study research: Combining rigour, relevance, and pragmatism," *Information Systems Journal*, vol. 8, pp. 273-289, 1998.
9. J. W. Creswell, *Research design: Qualitative, quantitative, and mixed methods approaches*, 2nd ed. Thousand Oaks, CA: Sage Publications, 2003.
10. M. A. Hjalmarson, "Learning from students’ responses to MEAs," in *Models and modeling in engineering education: Designing experiences for all students*, J. S. Zawojewski, H. Diefes-Dux, and K. Bowman, Eds. Rotterdam, The Netherlands: SensePublishers, 2008, pp. 173-186.
11. T. J. Moore and M. Hjalmarson, "Developing measures of roughness: Problem solving as a method to document student thinking in engineering " *International Journal of Engineering Education*, in press.
12. J. Wang, T. J. Moore, S. Plumb, and G. Roehrig, "A student task model method for assessing and improving a Model-Eliciting Activity," in *Frontiers in Education Conference*, San Antonio, TX, 2009.
13. N. Katims and R. Lesh, *PACKETS: A guidebook for inservice mathematics teacher development*. Lexington, MA: D.C. Heath, 1994.
14. S. M. Iversen and C. J. Larson, "Simple thinking using complex math vs. complex thinking using simple math – A study using Model Eliciting Activities to compare students’ abilities in standardized tests to their modeling abilities," *ZDM - The International Journal on Mathematics Education*, vol. 38, pp. 281-292, 2005.
15. S. Chamberlin and S. Moon, "Model-Eliciting Activities as a Tool to Develop and Identify Creatively Gifted Mathematicians," *Journal of Secondary Gifted Education*, vol. 17, pp. 37-47, 2005.
16. T. J. Moore, H. A. Diefes-Dux, and P. K. Imbrie, "Spontaneous groups versus long-term teams: An investigation using complex problem solving in a first-year engineering course " in *American Society for Engineering Education Annual Conference*, Honolulu, HI, 2007.
17. T. J. Moore, H. A. Diefes-Dux, and P. K. Imbrie, "Assessment of team effectiveness during complex mathematical modeling tasks," in *Frontiers in Education Conference*, San Diego, CA, 2006.

18. T. J. Moore, H. A. Diefes-Dux, and P. K. Imbrie, "The quality of solutions to open-ended problem solving activities and its relation to first-year student team effectiveness," in *American Society for Engineering Education Annual Conference*, Chicago, IL, 2006.
19. R. Lesh and H. M. Doerr, "Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching," Mahwah, NJ: Lawrence Erlbaum, 2003.
20. L. G. Militello and J. B. Hutton, "Applied Cognitive Task Analysis (ACTA): A practitioner's toolkit For understanding cognitive task demands," *Ergonomics*, vol. 41, pp. 1618-1641, 1998.
21. R. Lesh and S. J. Lamon, "Assessment of authentic performance in school mathematics," Washington, DC: AAAS Press, 1992.
22. G. Wiggins, *Educative assessment: Designing assessment to inform and improve student performance*. San Francisco: Jossey-Bass, 1998.
23. H. A. Diefes-Dux, T. J. Moore, J. Zawojewski, P. K. Imbrie, and D. Follman, "A framework for posing open-ended engineering problems: Model-Eliciting Activities," in *2004 Frontiers in Education Conference*, Savannah, GA, 2004.