AC 2011-1314: FEEDBACK AND ASSESSMENT OF STUDENT WORK ON MODEL-ELICITING ACTIVITIES: UNDERGRADUATE TEACHING ASSISTANTS' PERCEPTIONS AND STRATEGIES

Raghavi Merugureddy, Purdue University

Raghavi Merugureddy is a Senior in School of Industrial Engineering at Purdue University with minors in Electrical Engineering and Mathematics. Since 2007, she has been a member of Purdue's Society of Women Engineers (SWE), and Women in Engineering Program (WIEP). She has been a Vice President of Marketing for American Indian Foundation (AIF) chapter at Purdue from 2007 to 2008. Her research interest is on TA's assessment of student's open-ended solution on Model Eliciting Activities (MEAs).

Amani Salim, Purdue University, West Lafayette

Amani Salim is a postdoctoral researcher in the Department of Agricultural and Biological Engineering (ABE) at Purdue University, and was previously a postdoctoral researcher in the School of Engineering Education at Purdue University. She receives her B.Sc. and M.Sc. in Electrical Engineering from University of Minnesota Twin Cities, and her Ph.D. in BioMEMS and Microelectronics from Weldon School of Biomedical Engineering at Purdue University. Her engineering education research focuses on problem formulation within Model-Eliciting-Activities (MEAs) and engineering education, and professional development of teaching assistants.

Heidi A. Diefes-Dux, Purdue University, West Lafayette

Heidi 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 in Purdue's First-Year Engineering Program, the gateway for all first-year students entering the College of Engineering. She is currently the Director of Teacher Professional Development for the Institute for P-12 Engineering Research and Learning (INSPIRE). Her research interests center on implementation and assessment of mathematical modeling problems.

Monica E Cardella, Purdue University, West Lafayette

Monica E. Cardella is an Assistant Professor of Engineering Education and is the Co-Director of Assessment Research for the Institute for P-12 Engineering Research and Learning (INSPIRE) at Purdue University. Dr. Cardella earned a B.Sc. in Mathematics from the University of Puget Sound and an M.S. and Ph.D. in Industrial Engineering at the University of Washington. At the University of Washington she worked with the Center for Engineering Learning and Teaching (CELT) and the LIFE Center (Learning in Informal and Formal Environments). She was a CASEE Postdoctoral Engineering Education Researcher at the Center for Design Research at Stanford before beginning her appointment at Purdue. Her research interests include: learning in informal and out-of-school time settings, pre-college engineering education, design thinking, mathematical thinking, and assessment research.

Feedback and Assessment of Student Work on Model-Eliciting Activities: Undergraduate Teaching Assistants' Perceptions and Strategies

Abstract

Model-Eliciting-Activities (MEAs) are open-ended engineering problems that engage students in authentic modeling situations that professional engineers encounter. For seven years, Graduate Teaching Assistants (GTAs) were the primary point of contact and source of feedback for students during the implementation of MEAs in a large first-year engineering course. The recent addition and change in role of a significant number of Undergraduate Teaching Assistants (UGTAs) to the instructional team created the need for modifications to the TA Professional Development (PD) with MEAs. The purpose of this paper is to investigate UGTAs perceptions of and strategies for providing feedback and assessment of student work and summarize the challenges faced by them. Further studies will utilize these findings to modify the current *MEA Rubric* and task specific support materials and improve the TA PD.

I. Introduction

A. Open-Ended Problems and Undergraduate Teaching Assistants

One purpose of open-ended problem solving in engineering education is to emulate professional practice – provide students with authentic problem-solving experiences. Engagement in these experiences alone is not sufficient for meeting multiple learning objectives that may range from developing conceptual understandings to developing analytical skills to developing engineering "habits of mind"[1]. There must be supports for providing high impact instruction, feedback, and assessment that ensure learning along desired trajectories. These might include instructor training, mentoring, and task specific support materials. How these supports are used by instructors will depend on, among other things, their educational background, teaching experience, and professional experience[2,3]. The level of comfort and skill with which instructors can engage in teaching through open-ended problems will greatly affect the potential for student learning[4-6].

Due to large course enrollments, finances, and retention concerns, first-year programs often use undergraduate teaching assistants (UGTAs) (also known as peer teachers or peer learning assistants) to support classroom instruction, where their duties include providing classroom aid, functioning as liaison between students and faculty, preparing lesson plans, grading and tutoring[7,8]. Undergraduates have served as TAs at undergraduate institutions, where there is no graduate student pool to draw from, and in large entry-level courses[9,10]. At these undergraduate institutions, utilizing undergraduates as teaching assistants has shown to provide benefits – faculty spend more time improving their teaching, and undergraduate TAs themselves enrich their educational experience[7,11]. However, entry level science and mathematics courses tend to employ less complex ways of assessing learning outcomes than first-year courses involving open-ended problems with multiple solutions. Can UGTAs be used to support instruction with open-ended problem solving when complex learning outcomes are desired?

The evaluation of students' work on open-ended problems is challenging because multiple acceptable solutions exist[12]. This means that instructors need to be able to make sense of student work, evaluate its quality, and formulate feedback that will advance the students' thinking. It has been shown that graduate teaching assistants (GTAs) struggle to evaluate students' mathematical models across multiple dimensions[11]. Also, it was found that GTAs had difficulty evaluating students' problem formulations[13]. In both cases, the reliability of the GTAs' evaluations seemed linked to training and the availability of appropriate task specific supports. Some of the challenges in assessing student work were articulated in two case studies focused on GTAs' experiences[14]. Among these was included the challenge faced by GTAs in assessing students' responses when multiple solutions exist. Due to this, GTAs faced several conflicts: (1) balancing their roles as grader and guide, (2) dealing with progressive grading biases – providing consistent feedback across a number of pieces of student work, and (3) engaging in high quality feedback while addressing time management – providing high quality feedback is a time consuming process.

To begin to answer the question of whether UGTAs can support student learning through open-ended problem solving, we investigated UGTAs' experiences in assessing student work on open problems. Our objective was to understand their self-reported level of comfort and ability in assessing student team work along various dimensions. Additionally, we wanted to understand their approach to providing student teams with feedback and evaluation of their work. Further, we wanted to know to what degree they use various support materials provided to help them in their assessment task.

B. Model-Eliciting Activities (MEAs)

Model-Eliciting Activities (MEAs) are open-ended mathematical modeling problems set in engineering contexts[15]. The design of an MEA is based on six principles originally outlined by Lesh, et al.[16] and modified for engineering instruction[15]. A student team solution to an MEA is a generalizable (share-able, re-usable, modifiable) procedure (mathematical model) that can be used by a specified direct user to solve the given problem and similar problems. Student team solutions vary in both approach (with both multiple feasible and non-feasible approaches being put forth) and degree of development.

From Fall 2002 to Spring 2009, MEAs were implemented by GTAs in the laboratory setting of a required first-year engineering course at Purdue University. During this period, UGTAs were not involved in assessing student work on MEAs, though they did support classroom implementation. However, in Fall 2009, UGTAs, serving as either peer teachers (classroom instructional team members and graders) or out-of-classroom graders, became equally responsible with the GTAs for providing feedback on and evaluating students' MEA work. This recent staffing change brings challenges to implementing open-ended engineering problems. UGTAs, particularly sophomores, have minimal academic, teaching, or professional experience, as compared to GTAs, to rely on while grading or giving constructive feedback on student work. Further, they have severe time constraints for training and assessment of student work. This meant revising the established GTA professional development around MEAs[17] and making it accessible for UGTAs.

Since 2003, GTAs experiences, suggestions, and work products have been used to help develop the GTA professional development around MEAs and MEA evaluation tools. We have considered GTAs to be active participants with the faculty and researchers in the course reform and education of our students[17]. The change in the UGTAs role on the instructional team makes them, more than ever to us, what Seymour calls "partners in innovation"[18]. Their reflections on teaching through MEAs will likely lead to transformations in MEA implementation, TA professional development, TA mentoring, and MEA generic and task specific support materials - all to the benefit of students' learning through open-ended problems.

II. Research Questions

In this study, we examine UGTAs' experience with assessing student team work on MEAs. The evaluation tool used by all TAs is the four-dimension *MEA Rubric* which assesses the student teams' mathematical model and its generalizability (i.e. share-ability, re-usability and modifiability).

The research questions guiding this study are:

1) What are UGTAs' self-reported ability to apply the four dimensions of the *MEA Rubric*? What do UGTAs' report as being easy or challenging about using the four dimensions of the *MEA Rubric*.
2) What approaches do UGTAs use when providing feedback on student teams' MEA work?
3) Which MEA training and supports materials do UGTAs use to help themselves assess student teams' work? How frequently is each used?

III. Methods

A. Setting & MEA Implementation

The setting for this study was the Spring 2010 offering of a required first-year engineering course with an enrollment of approximately 1,300 students. This course was Part II of a two-semester sequence. In Part I, students were introduced to engineering problem solving, design, and teaming concepts. In Part II, engineering problem solving, design and teaming concepts were reinforced; computational tools and modeling were also introduced. Course meetings included two 110-minute periods per week led by a faculty member or GTA and supported by peer teachers. Each section of the course consisted of a maximum of 30 teams of four students.

In Spring 2010, two MEAs were implemented: *Travel Mode Choice* and *Nano Roughness*. For the *Travel Mode Choice* MEA, student teams must develop a model to predict students' modes of travel to campus at a growing university[19]. The *Nano Roughness* MEA is set in the context of manufacturing surface coatings for biomedical implants; student teams must design a procedure to quantify roughness from atomic force microscopy images of the surface coatings[19,20]. For each MEA, student teams are required to construct a share-able, re-usable, and modifiable mathematical model in the form of a written procedure (i.e. memo) intended for use by an identifiable direct user. MEA implementation involves two draft stages. Draft 1 enters a double-blind peer review. Draft 2 and the Final Team Response each receive feedback and evaluation from the TAs[21]. During the problem-solving process, students are provided data on which to base and test their models. Each MEA was implemented over a 6 week period.

B. TAs Roles with Regards to MEAs

Seven GTAs and 62 UGTAs (48 peer teachers and 14 out-of-class graders) were employed in Spring 2010. One GTA, four peer teachers, and one grader were assigned to each section of the course. Each UGTA was then typically responsible for providing feedback and assessment on the MEA work of five student teams. Each GTA was assigned to two sections and therefore was responsible for assessing the work of ten student teams.

All TAs received approximately 5 hours of professional development (PD) with each MEA. The PD for each MEA was completed in three phases:

- <u>Phase 1</u>: Prior to the PD session, TAs solved the MEA individually and then applied the four- dimension *MEA Rubric* to their work. The purpose of this phase was for TAs to become familiar with the MEA and the *Instructor's MEA Assessment/Evaluation Package* (I-MAP) and *MEA Feedback and Assessment Rubric* (*MEA Rubric* for short) before attending PD. An I-MAP provides MEA-specific information on how to apply the generic *MEA Rubric*[12].
- <u>Phase 2</u>: In a 2.5 hour face-to-face PD session, the course coordinator lead the TAs in a discussion of the role of open-ended problems and MEAs in a first-year engineering course and the TAs' role during MEA implementation and assessment. The bulk of the time was spent reviewing the I-MAP and *MEA Rubric* dimensions, allowing the TAs time to practice applying the rubric to two sample student team solutions, and then discussing their practice assessments.
- <u>Phase 3</u>: Following the PD session, TAs individually graded three prototypical pieces of student team MEA work using the online MEA system. This system allowed the TAs to assess each sample piece of student work, view an expert's assessment of that sample work, and reflect on how to improve their assessment.
- <u>Phase 4</u>: The GTA assigned to a given section reviewed his or her UGTAs' training assessments completed in Phase 3 and gave his or her UGTAs additional feedback on how to grade more effectively.

This PD differed from previous GTA-PD in a number of ways. The time spent in face-to-face training prior to the first MEA implementation was reduced from 8 to 2.5 hours. This was done by creating Phase 1 – requiring TAs to read and work the MEA before Phase 2. Completion of this phase was checked by the Lab Director. This was necessary as the first PD session had to be moved from before the semester started (when GTAs are required to be available for training) to an evening during the semester (when UGTAs are on campus). PD content associated more generally with teaching first-year engineering students had to be removed to shorten Phase 2. The number of prototypical pieces of student work the TAs practiced with in Phase 3 also had to be reduced from five to three due to time constraints. This limited the variety of potential student responses the TAs saw before they started to assess actual student work. In addition, individual feedback to TAs on their training (Phase 4) had been completed by a course instructor in previous semesters; the number of TAs made continuing this practice prohibitive. The GTAs were all experienced with MEAs (all with at least 3 semesters of teaching experience) and were coached on how to interpret the training results and address issues with their UGTA teams. The high level of experience of the GTAs was due to the retaining of our most experienced GTAs during the transition from using primarily GTAs to UGTAs.

C. *MEA Rubric*

UGTAs and GTAs conducted their feedback and assessment of student work by referencing the I-MAP and using the *MEA Rubric*. In Spring 2009, the *MEA Rubric* was updated from three[21] to four dimensions[22] as recommended in prior research[12]. These four dimensions are:

- Mathematical Model: Does the mathematical model adequately address the complexity of the problem?
- Re-usability: Can the direct user use the model on similar types of data?
- Modifiability: Can the direct user modify the model for use in similar but different situations?
- Share-ability: Can the direct user reproduce the results using the test case data provided in the MEA?

TAs applied the *MEA Rubric* to student teams' Draft 2 and Final Team Responses for each of the two MEAs implemented in Spring 2010.

D. Data Collection & Analysis

Participants

At the end of the Spring 2010 semester, all TAs were invited to participate in a survey conducted via a web-based interface. This survey was conducted to gain an understanding of their perceptions and practice regarding assessing student work on MEAs. Thirty-six TAs participated in this survey - 31 UGTAs and 5 GTAs or non-undergraduate students.

Data considered in this study are limited to responses received from UGTAs, of whom 28 were peer teachers and 3 were graders. The UGTA respondents consisted of 21 sophomores, 6 juniors, and 4 seniors as well as 16 males and 15 females. The academic year and gender distribution was representative of the entire UGTA population for the course. The high percentage of sophomores was due to the need to triple the number of UGTAs in AY 2009-10 as the course transitioned from using primarily GTAs to UGTAs. Recruitment of sophomores was given high priority due to their experience with the course and new facility in which the course was taught. The potential to train and retain these students as TAs for multiple semesters, even into their graduate programs, is high. But in addition, applications were distributed to current UGTAs for a version of the course that was being phased out, a peer mentoring program, a Women-In-Engineering distribution list of sophomores and juniors, and honors UGTA candidates who were not selected for the honors program, and professor recommended candidates. Word of mouth was also relied upon. Distribution to the Women-In-Engineering list may explain the equal number of male and female UGTAs. However, in semesters where the list was not used, the gender distribution of applications and hired UGTAs was still about equal. The appeal of the teaching task may be attracting a higher percentage of women than is representative of the undergraduate engineering population as a whole at Purdue University.

All those responding to the survey reported English to be their first language or they had at least 4 years of English background. Six out of the 31 UGTAs whose survey input was analyzed did not have an MEA experience in their first-year engineering course.

Overview of the Survey

The overall purpose of the survey was to understand the challenges TAs faced when using the *MEA Rubric* to provide feedback on and assess student responses to MEAs. The survey consists of 25 qualitative and quantitative items; some with multiple subparts (Appendix A). The survey is divided into 5 parts:

1. General Demographics – this captures gender, international status, and English proficiency
2. Teaching and MEA Experience – this includes TA classification, first-year experience with MEAs, and TA experience with MEAs
3. Experience Using the Rubric and Providing Feedback – items in the section relate to the TA's perceptions of their ability to apply each of the four dimensions of the *MEA Rubric*
4. Feedback Strategies – these items focus on how TAs give feedback and the resources they use
5. Value of MEAs and Feedback – includes questions about how TAs perceive the value of MEAs and feedback to student learning. Analysis of these items is not included in this paper.

Data Analysis

Subjects' quantitative responses to survey items 11 and 12 (ability to apply the four dimensions of the *MEA Rubric*) were analyzed using mean, standard deviation, and Pearson correlation analysis. Pearson correlation coefficient was used to test for significance, or strength, of correlations between two quantitative variables from the survey. The correlation is constrained between the values of -1 to 1 where the correlation value of 1 or -1 represents a "perfect linear relationship". A value of -1 represents an inverse relation between the variables[23]. The analysis was conducted with a significance level of 5 percent. More specifically, the coefficient was critiqued using the criteria in Table 1.

Table 1. Criteria used to critique Pearson correlation coefficient.

Criteria	**Indicates**
0	no linear relationship
+1	perfect positive linear relationship
-1	perfect negative linear relationship
between 0 and 0.3 (0 and -0.3)	weak positive (negative) linear relationship
between 0.3 and 0.7 (-0.3 and -0.7)	moderate positive (negative) linear relationship
between 0.7 and 1.0 (-0.7 and -1.0)	strong positive (negative) linear relationship

The frequency of response was computed for items 20 and 21 (feedback resources and strategies, respectively). Qualitative items were analyzed using open-coding to identify themes in participants' responses.

IV. Results

A. UGTAs' Self-Reported Ability to Apply the *MEA Rubric*

Survey items 11 and 12 asked UGTAs for their assessment of their comfort and ability to provide feedback for along each of the four *MEA Rubric* dimensions: Mathematical Model, Re-usability, Modifiability, and Share-ability. TAs considered five statements per dimension (Figure 1, Table 2, and Appendix A):

1. I was comfortable assigning marks in the _______ category of the rubric.
2. I was capable of providing a good critique in the ________ category of the rubric.
3. I could identify problems in student work related to the __________ category of the rubric.
4. I could provide written feedback related to the __________ category of the rubric.
5. I thought the teams I was reviewing would value the feedback I was providing in the ________ category of the rubric.

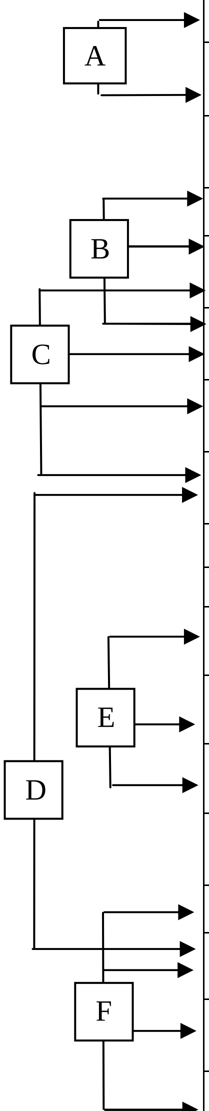
To respond to these items, the TAs were given the option of choosing: Strongly Agree, Agree, Neither Agree or Disagree, Disagree, Strongly Disagree. For analysis purposes, these options were assigned numerical values from 1 to 5, respectively. These numerical values were used to compute Pearson correlation coefficients which were then used to isolate the more problematic areas faced by the TAs in applying each dimension of the rubric.

Pearson correlation analysis in SAS resulted in a significant coefficient or strong positive relation between several subparts of items 11 and 12. Viewing the analysis results, it became apparent that these relations can be grouped into six clusters: A = [3, 4]; B = [6, 7, 8]; C = [7, 8, 9, 10]; D = [10, 17]; E = [12, 13, 14]; and F = [16, 17, 18, 19] as shown in Figure 1. These relations have a Pearson correlation coefficient of 0.7 or above, which according to the criteria in Table 2 represents a strong or significant relation. Pearson correlation coefficients greater than 0.9 where found for item pairs in cluster F (16&17), F (17&19), and F (18&19). Pearson correlation coefficients greater than 0.8 were found for item pairs in cluster A (3&4), B (6&7), C (9&10), D (2&14), E (12&13), E (13&14), F (16&19), and F (17&18). See Appendix B for more detailed results.

With regards to the Mathematical Model dimension of the *MEA Rubric*, cluster A shows that there is a strong dependency in the responses given by the UGTAs as to whether they could identify problems in student work and whether they could provide written feedback for this dimension.

For the Re-usability dimension, cluster B and C illustrate that the UGTAs' responses have a strong dependency among all five questions in this category. In other words, all the items regarding Re-usability have similar responses from the UGTAs. On the other hand, cluster D shows that the UGTAs had similar responses regarding whether they thought the team they were reviewing would value their feedback in Re-usability and whether they felt capable of providing good critique in the Share-ability sections of the category.

For the Modifiability dimension, cluster E shows UGTA's responses were similar in whether they felt capable of providing a good critique, whether they could identify problems in student work, and whether they could provide written feedback.



Q11	*When providing feedback on a student team solution:*
1	I was comfortable assigning marks in the MATHEMATICAL MODEL category of the rubric.
2	I was capable of providing a good critique in the MATHEMATICAL MODEL category of the rubric.
3	I could identify problems in student work related to the MATHEMATICAL MODEL category of the rubric.
4	I could provide written feedback related to the MATHEMATICAL MODEL category of the rubric.
5	I thought the teams I was reviewing would value the feedback I was providing in the MATHEMATICAL MODEL category of the rubric.
6	I was comfortable assigning marks in the RE-USABILITY category of the rubric.
7	I was capable of providing a good critique in the RE-USABILITY category of the rubric.
8	I could identify problems in student work related to the RE-USABILITY category of the rubric.
9	I could provide written feedback related to the RE-USABILITY category of the rubric.
10	I thought the teams I was reviewing would value the feedback I was providing in the RE-USABILITY category of the rubric.
Q12	*When providing feedback on a student team solution:*
11	I was comfortable assigning marks in the MODIFIABILITY category of the rubric.
12	I was capable of providing a good critique in the MODIFIABILIT category of the rubric.
13	I could identify problems in student work related to the MODIFIABILITY category of the rubric.
14	I could provide written feedback related to the MODIFIABILITY category of the rubric.
15	I thought the teams I was reviewing would value the feedback I was providing in the MODIFIABILITY category of the rubric.
16	I was comfortable assigning marks in the SHARE-ABILITY category of the rubric.
17	I was capable of providing a good critique in the SHARE-ABILITY category of the rubric.
18	I could identify problems in student work related to the SHARE-ABILITY category of the rubric.
19	I could provide written feedback related to the SHARE-ABILITY category of the rubric.
20	I thought the teams I was reviewing would value the feedback I was providing in the SHARE-ABILITY category of the rubric.

Figure 1. Clusters of subparts of items 11 and 12 with a Pearson correlation coefficient of 0.7 or greater.

Cluster F shows that most of the subparts related to the Share-ability dimension have responses similar to each other except for whether the UGTA thought the feedback they provided was valued.

In summary, these clusters in Figure 1 that do have similar responses can be viewed as a single item with a single response. Table 3 shows the average response for individual items and clusters resulting from the Pearson analysis and a qualitative representation of the numeric mean for these items and clusters (where 1 to 1.5 was assigned Strongly Agreed, 1.5 to 2.5 Agreed, and 2.5 to 3.5 Neutral).

As seen in Table 3, UGTAs self-reported strongly agreeing that they were comfortable and capable of providing feedback and assessing student work along the Share-ability dimension, and they felt their feedback would be valued by the students. They also self-reported agreeing to these same things for the Re-usability dimension. To a lesser extent, they agreed to these for the Mathematical Model dimension. For the Modifiability dimension, the UGTAs were, on average, slightly closer to neutral in their self-reported comfort at assigning marks. Across the Mathematical Model and Modifiability dimension items, there was greater variability in the UGTA responses, indicating that these dimensions were more difficult for some UGTAs to apply.

B. UGTAs' Strategies for Providing Feedback

UGTAs responses regarding how often they employed particular strategies for providing feedback are shown in Table 4. The majority of UGTAs agreed that they always or frequently "read over the whole piece of work" before they made any comments (item 21a). This corresponded to their reporting of only occasionally writing comments as they read through a piece of student work (item 21h). UGTAs always or frequently were "aware of what feedback students have and have not responded to" (item 21b). Though to a much lesser extent the UGTAs always or frequently reviewed TA feedback to the team on Draft 2 (item 21e) and peer feedback on Draft 1 (item 21g). UGTAs always or frequently tried to use students' procedures to generate results (item 21d); though to a lesser extent they made notes about the mathematics students were using (item 21f). The UGTAs report that they wrote their feedback directly through the MEA online interface for one piece of student work at a (item 21c) as opposed to offline and uploading them one at a time or individually (items 21i and 21k, respectively). Only occasionally did the UGTAs report going back and regrading a piece of student work (item 21j).

C. UGTAs' Use of Training and Support Materials

UGTAs responses regarding what materials they used to help provide feedback are shown in Table 5. The "MEA IMAP – Grading Guide" (item 20a) and "MEA Rubric online" (item 20c) were the materials that UGTAs referenced the most - 65% and 45% of UGTAs, respectively, reported they "always have this open or available" when providing feedback. UGTAs also tended to refer, at least periodically, to the data sets included in the MEAs (item 20b), MEA instructions (item 20d), and online expert feedback (item 20e). The material that UGTAs referenced the least when providing feedback was "the solutions to the MEA that they submitted prior to the training workshop" (item 20h) – 36% said they referenced the material only once and another 48% did not find the material useful as a reference.

Table 3. UGTAs' responses modified based on results from Pearson correlation analysis.

***MEA Rubric* Dimension**	**Summarized Questionnaire after Pearson Analysis**	**Mean**	**Numerical Representation**
Item 11	*When providing feedback on a student team solution:*		
Mathematical Model	1. I was comfortable assigning marks in the MATHEMATICAL MODEL category of the rubric.	2.03	Agreed
	2. I was capable of providing a good critique in the MATHEMATICAL MODEL category of the rubric.	1.81	Agreed
	CLUSTER A: 3 & 4. I could identify problems in student work and provide written feedback related to the MATHEMATICAL MODEL category of the rubric.	1.71	Agreed
	5. I thought the teams I was reviewing would value the feedback I was providing in the MATHEMATICAL MODEL category of the rubric.	1.87	Agreed
Re-Usability	CLUSTER B: 6, 7 & 8. I was comfortable assigning marks, capable of providing a good critique, and could identify problems in student work in the RE-USABILITY category of the rubric.	1.78	Agreed
	CLUSTER C: 7, 8, 9, & 10. I was capable of providing a good critique, could identify problems in student work, could provide written feedback, and thought to value the feedback provided in the RE-USABILITY category of the rubric.	1.77	Agreed
Item 12	*When providing feedback on a student team solution:*		
Modifiability	11. I was comfortable assigning marks in the MODIFIABILITY category of the rubric.	2.42	Agreed
	CLUSTER E: 12, 13 & 14. I was capable of providing a good critique, could identify problems in student work, and could provide written feedback in the MODIFIABILITY category of the rubric.	2.00	Agreed
	15. I thought the teams I was reviewing would value the feedback I was providing in the MODIFIABILITY category of the rubric.	2.06	Agreed
Share-ability	CLUSTER F: 16, 17, 18, & 19. I was comfortable assigning marks, capable of providing a good critique, could identify problems in the student work, and could provide written feedback in the SHARE-ABILITY category of the rubric.	1.51	Strongly Agreed
	20. I thought the teams I was reviewing would value the feedback I was providing in the SHARE-ABILITY category of the rubric.	1.58	Strongly Agreed

Table 4. UGTAs responses towards how often they take various actions when providing feedback on student work on MEAs (n = 31).

21. When providing feedback on a student team's solution to an MEA…	Always	Frequently	Occasionally	Never
a. I read over the whole piece of work before I make any comments	58%	26%	13%	3%
b. I am aware of what feedback students have and have not responded to	48%	42%	10%	0%
c. I write my feedback directly through the MEA online interface	52%	23%	22%	3%
d. I use the students' procedure (work through it to see what results I get)	39%	52%	6%	3%
e. When grading MEA Team Final, I review my (or previous TA) feedback to the team on Draft 2	45%	29%	19%	6%
f. I make a note of the mathematics that students are using	29%	48%	19%	3%
g. When grading MEA Draft 2, I review the peer feedback to the team on Draft 1	16%	35%	26%	23%
h. I write down comments as I am reading the student team work	16%	29%	48%	6%
i. I write my feedback offline for all student work before uploading and submitting	16%	10%	6%	68%
j. I go back and regrade after reading more student work	10%	23%	45%	23%
k. I write my feedback offline for one piece of student work at a time, upload and submit, before going on to another	3%	0%	32%	65%

Table 5. UGTAs responses towards how often they referenced various support materials when providing feedback on student work on MEAs (n = 31).

20. I use the following materials to help me provide feedback:	I always have this open or available	I reference this periodically	I reference this once	I never use this because…	
				I did not find it useful	I forgot it is available
a. MEA IMAP - Grading Guide	65%	23%	10%	0%	3%
b. Data sets for the MEA	39%	48%	10%	3%	0%
c. MEA Rubric online	45%	26%	16%	3%	10%
d. MEA class or homework instructions	23%	39%	26%	6%	6%
e. Expert feedback from online training	23%	39%	19%	19%	0%
f. Feedback I submitted during online training	16%	29%	19%	32%	3%
g. PowerPoint slides from MEA workshops with Prof. X	13%	13%	26%	23%	26%
h. The solution to the MEA that I submitted prior to the training workshop	6%	3%	35%	48%	6%

V. Discussion

According to the quantitative results presented above and UGTAs response to item 13, the dimensions of the *MEA Rubric* that the UGTAs found difficult or challenging to provide feedback on were Mathematical Model (45%) and Modifiability (42%). This is consistent with findings for GTAs when studying their reliability at applying an earlier version of the *MEA Rubric*[12]. In response to why these dimensions were challenging and how they worked around these difficulties, 23% of the UGTA respondents wrote about how difficult it was to interpret students' mathematical models. One UGTA responded:

> *"It was the most difficult to grade the Mathematical model because every group composed a completely different idea. If the students were not proficient at explaining the mathematical model, then it was extremely difficult to follow the rubric. A lot of the math is usually implied and the students take that for granted."*

In their written responses, nineteen percent stated that for the Mathematical Model dimension, they found the criteria to not be as clear for the second MEA implemented in Spring 2010. Here is a sample of a UGTA's response:

> *"A lot of the times it was unclear what exactly qualified in each of the categories of the Math Model, especially in the second MEA. In the first MEA, I was completely fine and understood everything, but in the second, I had a hard time determining what I should take off for or even count as good. Sometimes the expert seemed a little inconsistent in the second MEA, so it was hard trying to determine what was okay and what wasn't in the second MEA. I also had a hard time understanding what the students were trying to do with regard to their procedure with the gold data because they would use statistics, and my brain never really understood statistics that well. Thus, I had a hard time understanding what they wanted me to do and why they were doing certain things. It made giving them feedback really hard. I spent some time discussing my problems with my graduate TA, and used the powerpoint for the training religiously to help me get through it."*

UGTAs responded on item 14 that the Re-usability (51%) and Share-ability (35%) dimensions were the least difficult or challenging dimensions of the rubric to apply. For their reasons of why they felt this way, 48% of the UGTAs stated that Re-usability was very clear-cut, the criteria was clearly stated in the rubric, and having practice prior to grading the students' MEA solutions was also very helpful. Here is a sample UGTA's response:

> *"Re-usability was easy because they either correctly identified all of the necessary parts to this section or they didn't. It was very clear cut."*

UGTAs also agreed that the Share-ability dimension was very straight forward and expectations were clearly stated in the rubric. Therefore, there was very little to question as to what was expected of student work. A typical UGTA's response was:

> *"They either had the results and they were accurate. It is pretty cut and dry. There wasn't room for guessing or students kind of understanding, it was either right or wrong."*

These results are expected. TAs are provided with guidelines in the I-MAP for what mathematical ideas need to be embedded in a high quality mathematical model. Their responsibility is to interpret what mathematics students have used in their models and give feedback on the students' existing models in such a way as to guide them towards a higher quality model. TAs are also responsible for looking at how students justify the mathematics they use in their models. This all takes concentration, practice, and experience. Further, it is evident during TA training that many TAs do not know what it means to justify a decision made in a model, making assessing Modifiability a challenge. And since the justifications students need to make are dependent on the way their models are constructed, providing a list of reasonable justifications a priori in the I-MAP is difficult. Additional TA training with Modifiability needs to be considered; as does the provision of additional support for TAs with weak mathematical conceptual understandings.

TAs find Re-usability and Share-ability less challenging because explicit lists of things that must be included or excluded in student work is provided in the I-MAP. For instance, for Re-usability, the students must explain who the direct user is; what this user needs in terms of a deliverable and its criteria for success and constraints; and an overarching description of the model. For Share-ability, TAs are looking for results of the students applying their own model in a specified format, a model written in a way that can be replicated, and no extraneous information.

From Table 4, it can be deduced that a majority of UGTAs always read over a whole piece of work before making comments and they are always or frequently aware of what feedback to which students have and have not responded. This statement coincides with UGTAs' most popular qualitative response - 22% stated that their strategy for grading Draft 2 versus Team Final was to grade a team's final solution based on how they implemented the feedback they received for Draft 2. Here is a sample UGTA's response:

> *"For the team final draft, I focused more on if the students followed the feedback which was presented for Draft 2."*

This tendency of the UGTAs to assess student work based on whether they responded to feedback does give one pause. What constitutes a high quality model according to an MEA's I-MAP holds regardless of the solution iteration being completed and assessed. If TAs provide feedback and assessment based primarily on improvement, rather than the criteria for a high quality model, students may be left with an impression that they have achieved a high quality model. This strategy also leads to grade inflation, a problem that has begun to be noted in other parts of the larger study on feedback and assessment.

VI. Conclusions and Future Directions

In this study, UGTAs perceptions of their ability to apply and their strategies for applying a rubric to students' open-ended responses to mathematical modeling problems were investigated. It was found that UGTAs have more difficulty with the dimensions of the rubric that require interpretation of student work than those where check-lists could be applied. UGTAs read through and then provide online feedback on each piece of student work in succession, only occasionally going back to regrade. UGTAs most frequently use MEA I-MAP grading guides, the data sets included in the MEAs, the MEA online rubric, and expert feedback from online training to help themselves conduct feedback and assessment.

The results of the study indicate that future PD needs to better address feedback and assessment of students' mathematical models and how they are justified. PD also needs to emphasize that the criteria for a high quality model are fixed and apply across all solution iterations, even the final team response. Added support materials should be included in materials that TAs already most frequently use.

For others using or consider using UGTAs to support open-ended problem-solving, the results of this study provide some insight. UGTAs have confidence in their ability to apply rubrics that are check lists. That is, they are confident in their ability to conduct search-and-find assessment tasks in solutions to open-ended problem. UGTAs are less confident in their ability to interpret, provide feedback on, and assess students' solutions to open-ended problem given only guidelines for what constitutes a high quality solution. So, UGTAs need significant training and ongoing support for the latter.

Future research should look at the actual feedback and assessment UGTAs provide to students and its impact on students' thinking as revealed through their models. Further, as changes are made to TA PD with MEAs (or other open-ended problems), monitoring should be done to understand the impact of the PD on (1) TAs - their confidence in their feedback and assessment abilities, their feedback and assessment strategies, and their use of resources, and (2) students – the quality of their work products and ultimately their learning.

Acknowledgements

This work was made possible by grants from the National Science Foundation (DUE 0717508 and EEC 0835873). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Bibliography

1. Katehi, L., Pearson, G., & Feder, M. (2009). The status and nature of K–12 engineering education in the United States. *The Bridge*, *39*(3), 5-10.
2. Goodland, S. (1997). Responding to perceived training needs for graduate teaching assistants. *Studies in Higher Education, 22*(1), 83-92.
3. Salinas, M.F., Kozuh, G., & Seraphine, A. (1999). I think I can: Improving teaching self-confidence of international teaching assistants. *Journal of Graduate Teaching Assistant Development, 6*(3), 149-156.
4. Fink, L.D. & Ambrose, S. (2005). Becoming a professional engineering educator: A new role for a new era. *Journal of Engineering Education, 94*(1), 185-194.
5. Shulman, L. (1986). Paradigms and research programs in the study of teaching; A contemporary perspective. In M. C. Witrock (Ed.), *Handbook of Research in Teaching, 3rd ed.*, New York: Macmillan.
6. Bransford, J.D., Brown, A.L., and Cocking, R.R. (Eds.). (1999). Chapter 7: Effective Teaching: Examples in History, Mathematics, and Science. In *How People Learn, Brain, Mind, Experience, and School* (pp. 143-177). Washington D.C.: National Academy Press.
7. Jabker, E.H. & Reeves, S.G. (1976). Institution-wide use of Undergraduate Teaching Assistants. *Innovative Higher Education, 1*(1): 68-80.
8. Romm, I., Gordon-Messer, S., & Kosinski-Collins, M. (2010). Educating young educators: A pedagogical internship for undergraduate teaching assistants. *CBE- Life Sciences Education*, *9*(2), 80-86.

9. Vattano, F.J., Hockenberry, C., Grider, W., Jacobson, L., & Hamilton, S. (March 1973). Employing undergraduate students in the teaching of psychology. *APA Newsletter: Division on the Teaching of Psychology*, 9-13.
10. Firmin, M.W. (2008). Utilizing undergraduate assistants in general education courses. *Contemporary Issues In Education Research*, *1*(1), 1-6.
11. Osborne, R.E., Norman, J. & Basford, T. (1997). Utilizing undergraduate teaching assistants: An untapped resource. *Annual Conference on Undergraduate Teaching of Psychology,* Ellenville, NY.
12. Diefes-Dux, H.A., Zawojewski, J.S., & Hjalmarson, M.A. (2010). Using educational research in the design of evaluation tools for open-ended problem. *International Journal of Engineering Education*, *26*(4): 807-819.
13. Salim, A. & Diefes-Dux, H.A. (2010). Graduate teaching assistants' assessment of students' problem formulation within Model-Eliciting-Activities. *Proceedings of the American Society for Engineering Education (ASEE) Annual Conference*, Louisville, KY.
14. Cardella, M.E., Diefes-Dux, H.A., Verleger, M.A., & Oliver, A (2009). Insights into the process of providing feedback to students on open-ended problems. *Proceedings of the American Society for Engineering Education (ASEE) Annual Conference*, Austin, TX.
15. Diefes-Dux, H., Hjalmarson, M. A., Miller, T., & Lesh, R. (2008). Chapter 2: Model-Eliciting Activities for Engineering Education. In J.S. Zawojewski, H.A. Diefes-Dux, & K.J. Bowman (Eds.), *Models and Modeling in Engineering Education: Designing Experiences for All Students* (pp. 17-36). Rotterdam, The Netherlands: Sense Publishers.
16. Lesh, R., Hoover, M., Hole, B., Kelly, A., & Post, T. (2000). Principles for developing thought-revealing activities for students and teachers. In A.E. Kelly and R.A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 591-646). Mahwah, NJ: Lawrence Erlbaum.
17. Diefes-Dux, H.A., Osburn, K., Capobianco, B.M., & Wood, T. (2008). On the front line – Learning from the teaching assistants. In J.S. Zawojewski, H.A. Diefes-Dux, & K.J. Bowman (Eds.), *Models and Modeling in Engineering Education: Designing Experiences for All Students* (pp. 225-255). Rotterdam, The Netherlands: Sense Publishers.
18. Seymour, E., Melton, G., Wiese, D.J., & Pedersen-Gallegos, L. (2009). The importance of teaching assistants to undergraduates in the sciences, In *Partners in Innovation: Teaching Assistants in College Science Courses,* (pp. 7-14). Lanham, Maryland: Rowman and Littlefield Publishers.
19. Zawojewski, J.S., H.A. Diefes-Dux, & K.J. Bowman, (Eds). (2008). *Models and Modeling in Engineering Education*. Rotterdam, The Netherlands: Sense Publishers.
20. Moore, T.J. & Hjalmarson, M.A. (2010). Developing measures of roughness: problem solving as a method to document student thinking in engineering. *International Journal of Engineering Education*, *26*(4), 820-830.
21. Verleger, M., Diefes-Dux, H., Ohland, M.W., Besterfield-Sacre, M., and Brophy, S. (2010). Challenges to informed peer review matching algorithms, *Journal of Engineering Education*, *99*(4): 397-408.
22. Carnes, M.T., Diefes-Dux, H.A., & Cardella, M.E. (2011). Evaluating student responses to open-ended problems involving iterative solution development in Model-Eliciting Activities, *Proceedings of the American Society for Engineering Education (ASEE) Annual Conference*, Vancouver, CAN.
23. Moore, D.S. & McCabe, G.P. (2003). *Introduction to the Practice of Statistics* (pp. 126-130). 4th ed. New York: W.H. Freeman.

APPENDIX A: ONLINE TA SURVEY

Part 1: General Demographics

1. Name ________.
2. I am a *male / female*.
3. I am a(n) *international / domestic* student.
4. English is my first language. *Yes/ No*.
5. In my pre-college education, English was the primary language used for instruction. *Yes / No*
6. Is there anything else you would like to share about your English language background?

Part 2: Teaching and MEA Experience

7. I am a: *Sophomore / Junior / Senior / M.S. Student / Ph.D. Student*
8. I am a: *Peer Teacher / Grader / Graduate Teaching Assistant / Other*
9. I experienced MEAs as a first-year engineering student at Purdue University: *Yes / No*
10. I have X *semesters* of experience giving feedback to students on Model-Eliciting Activities (including this semester): _______

Part 3: Experiences Using the Rubric and Providing Feedback

To review the MEA rubric, right click on the following link to open it in a new window:
https://engineering.purdue.edu/ENE/Research/SGMM/Rubric_2009.pdf

11. Consider how much you agree or disagree with each of the following statements.

    As I completed providing feedback and assessment of first-year engineering student work on MEAs this semester:

	Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strong Disagree*
I was comfortable assigning marks in the **Mathematical Model** category of the rubric					
I was capable of providing a good critique in the **Mathematical Model** category of the rubric					
I could identify problems in student work related to the **Mathematical Model** category of the rubric					
I could provide written feedback related to the **Mathematical Model** category of the rubric					
I thought the teams I was reviewing would value the feedback I was providing in the **Mathematical Model** category of the rubric.					

	Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strong Disagree*
I was comfortable assigning marks in the **Re-Usability** category of the rubric					
I was capable of providing a good critique in the **Re-Usability** category of the rubric					
I could identify problems in student work related to the **Re-Usability** category of the rubric					
I could provide written feedback related to the **Re-Usability** category of the rubric					
I thought the teams I was reviewing would value the feedback I was providing in the **Re-Usability** category of the rubric.					

12. Consider how much you agree or disagree with each of the following statements.

    As I completed providing feedback and assessment of first-year engineering student work on MEAs this semester:

	Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strong Disagree*
I was comfortable assigning marks in the **Modifiability** category of the rubric					
I was capable of providing a good critique in the **Modifiability** category of the rubric					
I could identify problems in student work related to the **Modifiability** category of the rubric					
I could provide written feedback related to the **Modifiability** category of the rubric					
I thought the teams I was reviewing would value the feedback I was providing in the **Modifiability** category of the rubric.					
I was comfortable assigning marks in the **Share-ability** category of the rubric					
I was capable of providing a good critique in the **Share-ability** category of the rubric					
I could identify problems in student work related to the **Share-abilit**y category of the rubric					
I could provide written feedback related to the **Share-ability** category of the rubric					
I thought the teams I was reviewing would value the feedback I was providing in the **Share-ability** category of the rubric.					

13. Which dimension of the rubric was most difficult or challenging for you to provide feedback on? *Mathematical Model / Re-Usability / Modifiability / Shareability*

    Why? How did you work around these difficulties or challenges?

14. Which dimension of the rubric was least difficult or challenging for to provide feedback on? *Mathematical Model / Re-Usability / Modifiability / Shareability*

    Why? What made this dimension easier to provide feedback on?

15. Were there aspects of the rubric that were particularly useful or helpful in providing feedback?

16. What would help you give better feedback to students?

Part 4: Feedback Strategies

17. How long in minutes do you typically spend proving feedback on each piece of student team work?

18. How many student team solutions to an MEA do you usually evaluate at a time (in one sitting)? *1-10*

19 How do you select the order in which to assess the student team work you've been assigned to provide feedback on?

20. I use the following materials to help me provide feedback:

    I always have this open or available /
    I reference this periodically /
    I reference this once /
    I never use this because I did not find it useful
    I never use this because I forget it is available

    a. MEA class or homework instructions
    b. Data sets for the MEA
    c. MEA IMAP – Grading Guide
    d. MEA Rubric online
    e. Feedback I submitted during online training
    f. The solution to the MEA that I submitted prior to the training workshop
    g. PowerPoint slides from MEA workshops with Prof. X
    h. Expert feedback from online training

21. When proving feedback on a student team's solution to an MEA …

    Always / Frequently / Occasionally / Never

    a. I read over the whole piece of student work before I make any comments
    b. I write down comments as I am reading the student team work
    c. I make a note of the mathematics that students are using
    d. I use the students' procedure (work through it to see what results I get)
    e. I am aware of what feedback students have and have not responded to
    f. I go back and regrade after reading more student work
    g. I write my feedback directly through the MEA online interface
    h. I write my feedback offline for one piece of student work at a time, upload and submit, before going on to another
    i. I write my feedback offline for all student work before uploading and submitting
    j. When grading MEA Draft 2, I review the peer feedback to the team on Draft 1
    k. When grading MEA Team Final, I review my (or previous TA) feedback to the team on Draft 2

22. Was your strategy for providing feedback different for Draft 2s versus Team Finals? *Yes / No*

    Why or why not?

23. Did your strategy for providing feedback change from MEA 1 to MEA 2? *Yes / No*

    Why or why not? What was working or not working for you?

Part 5: Value of MEAs and Feedback

24. What do you think is the value of MEAs?
25. What do you want students to do or get as a result of receiving the feedback that you give on an MEA?

APPENDIX B: PEARSON CORRELATION ANALYSIS RESULTS

Table B1. Pearson correlation analysis - paired item results.

Item Pairs	Pearson Correlation	Significance p-values
Greater than 0.7		
6 & 8	0.75755	<0.0001
7 & 8	0.73404	<0.0001
7 & 9	0.71796	<0.0001
8 & 10	0.70717	<0.0001
10 & 17	0.72220	<0.0001
16 & 18	0.79570	<0.0001
Greater than 0.8		
3 & 4	0.81949	<0.0001
6 & 7	0.81166	<0.0001
9 & 10	0.80259	<0.0001
12 & 13	0.82789	<0.0001
12 & 14	0.84730	<0.0001
13 & 14	0.89808	<0.0001
16 & 19	0.86382	<0.0001
17 & 18	0.88411	<0.0001
Greater than 0.9		
16 & 17	0.93069	<0.0001
17 & 19	0.93773	<0.0001
18 & 19	0.90747	<0.0001