AC 2012-3820: THE NATURE OF PEER FEEDBACK FROM FIRST-YEAR ENGINEERING STUDENTS ON OPEN-ENDED MATHEMATICAL MODELING PROBLEMS

Miss Kelsey Joy Rodgers, Purdue University

Kelsey Rodgers is a graduate student at Purdue University in the School of Engineering Education. She is currently conducting research on peer feedback within model-eliciting activities (MEAs) in the First-year Engineering program with her advisor, Professor Heidi Diefes-Dux. Prior to attending Purdue, she graduated from Arizona State University with her B.S.E in engineering from the College of Technology and Innovation. She began her research in engineering education on disassemble, analyze, assemble (DAA) activities with her previous advisor at Arizona State University, Professor Odesma Dalrymple.

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

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. She is a member of Purdue's Teaching Academy. 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 has coordinated and taught in a required first-year engineering course that 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).

Dr. Monica E. Cardella, Purdue University, West Lafayette

The Nature of Peer Feedback from First-Year Engineering Students on Open-Ended Mathematical Modeling Problems

I. Introduction

Feedback is the process of identifying a gap between current and optimal solutions; then determining methods to advance the current work.[1,2] This process can be completed in a formal and informal manner. These two types of feedback processes are important elements of professional practice in the science, technology, engineering, and mathematics (STEM) community, including education, industry, and research.[3] An ability to provide feedback is also a fundamental skill that needs to be developed by students in STEM disciplines.[4,5]

In STEM education, effective teacher feedback is largely acknowledged to be one of the most important factors in student success. [6]It is also acknowledged to be one of the most lacking areas in quality and research.[6] In industry, employers and managers are unable to give complete feedback that satisfies employees need to understand their current successes and shortcomings.[7] In the STEM research community, fellow researchers must give feedback on most papers prior to publication, which means the STEM community could not disseminate research without effective feedback.[8] Ensuring that our STEM educators, professionals, and researchers are capable of effective feedback is crucial for the continuing success of the STEM community.

Due to the need for informal and formal feedback in STEM careers, it is vital to teach future STEM professionals how to develop and give effective feedback during their undergraduate education. Teaching peer feedback enables students to do more than just give effective feedback, it also enhances their communication, teaming, critical thinking, and problem solving skills, which are nationally and globally recognized key skills for engineering students to develop.[4,5]

Since feedback is a crucial element in STEM education, it is important for STEM education researchers and instructors to establish an efficient and effective pedagogy to teach STEM students how to give peer feedback to improve others work and receive feedback to advance their own work. To create effective pedagogies to teach STEM undergraduates how to give high-quality, effective peer feedback, the current quality and nature of STEM students' peer feedback must be established.[9]

II. Background

This research on peer feedback is being conducted within the Model-Eliciting Activities (MEAs) implementation sequence (Figure 1). MEAs are open-ended modeling problems that challenge students to work in teams to solve complex problems with realistic applications.[10] The implementation strategy of these activities involves multiple iterative points that challenge students to constantly and critically evaluate their team's solution and improve their solution using external feedback. The solutions are submitted in the form of a memo in which students communicate their developed procedure (or mathematical model or method for solving the problem), their mathematical logic behind their procedure (or why they selected their method for solving the problem), the potential limitations of their procedure, and the relevant assumptions made in their procedure.

Peer feedback is one vital step in the iterative process of the MEA implementation sequence (Figure 1). Previous research on students' perceptions of their ability to give peer feedback on MEAs shows that students believe they give good critiques of peers' work.[11] However, in an analysis of student team work on the *Just-In-Time (JIT) Manufacturing* MEA, it was revealed that major changes in students' mathematical models were made by only 11 out of 50 teams' solutions (22%) following peer feedback.[12] Since students think they are capable of giving good feedback and the results show a lack of improvement in teams' solutions after peer feedback, the focus of this study is on the quality and nature of peer feedback during an implementation of the *JIT Manufacturing* MEA. The three research questions that guide this work are: (1) What is the level of student participation in giving feedback? (2) How do peer reviewers scores compare to expert scores on student team MEA solutions? and (3) What aspects of students' solutions do students focus on in peer feedback?.

III. Research Context

A. Setting and Participants

In the first-year engineering program at Purdue University, all engineering students (1200-1700) are given the opportunity to learn how to solve authentic engineering problems in teams of 3 to 4 students through their engagement in MEAs.[13] In the semester of this study, Fall 2008, approximately 1200 students completed three MEAs. All MEAs are solved using the same implementation sequence (described in Figure 1) and assessed using the same *MEA Feedback and Assessment Rubric* (*MEA Rubric,* APPENDIX A).

This study is concerned with the second of these: the *JIT Manufacturing* MEA. For this MEA, teams develop a process to rank prospective shipping companies in order of best to least able to meet the client's timing needs based on historical late delivery time data for eight potential shipping companies. The full historical data includes 270 arrival times in minutes late, ranging from 0 to 100 minutes. All eight shipping companies have varying data with different distributions that have similar means (9.45 min. +/- 0.7 min.). The team's analysis of the data should help them understand the need to account for distribution of the data rather than just mean and/or standard deviation when developing their mathematical model.

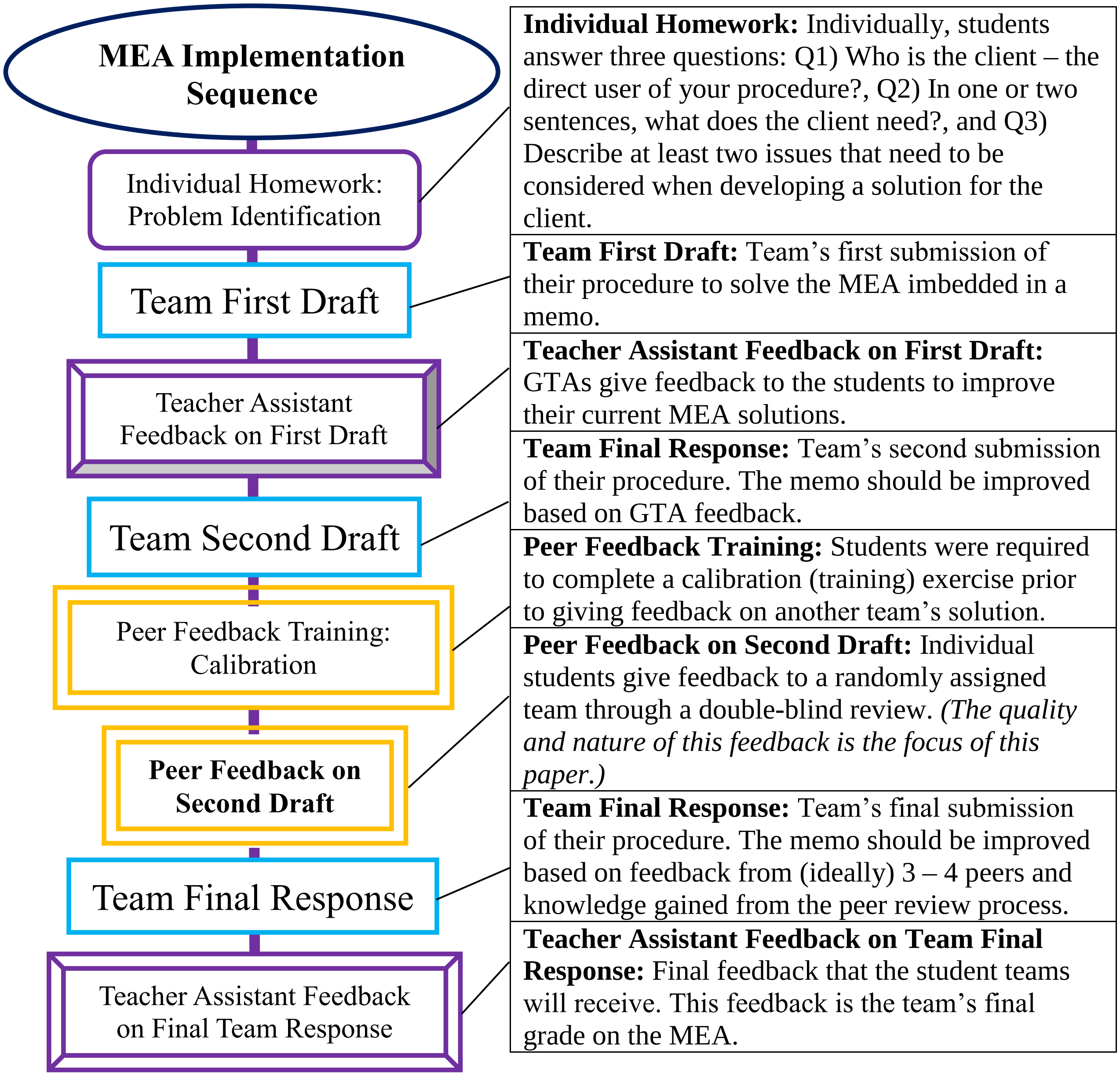

Figure 1. MEA Implementation Sequence for Fall 2008.

The peer calibration gives students access to a prototypical piece of student team work on the MEA under consideration through a web-based interface. Students individually critique the solution using the *MEA Rubric*. The students are then shown an expert critique of that team solution side-by-side with their critique. The students are prompted to compare and contrast their work to the expert work to improve their feedback abilities. Each student is then tasked to complete a double-blind peer review of one selected peer team. In this semester, the web-based interface that enabled the double-blind review assigned either peers to teams randomly or based on graduate teaching assistant (GTA)-assigned scores on the First Draft (Draft 1) as described by Verleger, et al.[14].

The GTA feedback, peer feedback, and final assessment of team solutions are all conducted using the *MEA Rubric*, which focuses on three dimensions: Mathematical Model, Re-Usability & Modifiability, and Share-Ability (Audience). The development of these dimensions is discussed by Diefes-Dux, Zawojewski, and Hjalmarson.[15] The Mathematical Model focuses on two aspects, complexity and rationales. Mathematical Model Complexity addresses whether or not the math utilized addresses the complexity embedded in the situated problem. Mathematical Model Rationale focuses on the team's ability to justify their reasoning behind their mathematical model. The Re-Usability & Modifiability dimension focuses on whether or not the solution can be used by the client for new but similar situations (e.g. additional historical data) and whether or not the solutions can be modified easily by the client for slightly different situations (e.g. ranking internet providers for a company based on speed and reliability instead of shipping companies on delivery times). The Share-Ability (Audience) dimension addresses whether or not it is easy for the client to follow and apply the solution.

B. Data Collection

The implementation of the MEA was supported by a web-based interface connected to a database system. This system was utilized to collect student work (e.g. individual assignments and team iterative solutions), manage GTA feedback and assessment, conduct peer calibration, and enable the double-blind peer review. To address the research questions, the focus here is on the peer feedback given to teams on their Second Draft solutions for the *JIT Manufacturing* MEA. Peer feedback consisted of numeric scores and written feedback for each of the three *MEA Rubric* dimensions.

Peer review required peers score eight items related to the three dimensions of the *MEA Rubric* (APPENDIX A). The Mathematical Model dimension has 4 sub-dimensions related to the existence of a model, a team solution's ability to address the complexity of the problem, the use of the given data, and the rationales that explain the reason for the selected procedure. The Re-Usability & Modifiability dimension is scored based on stated assumptions made in the solution. The Share-Ability dimension has 3 sub-dimensions related to the presentation of results, readability of the solution, and exclusion of extraneous information. The possible scores of each sub-dimension within the *MEA Rubric* are different but range from a minimum of 0 to a maximum of 4. The score ranges are different due to the fact that most sub-dimensions do not matter unless the mathematical model sufficiently addresses the complexity of the problem. The overall grade on a MEA team solution is the lowest score received on any *MEA Rubric* sub-dimension.

The students were also required to give written feedback in response to eight prompts associated with the three *MEA Rubric* dimensions (APPENDIX B). The written feedback was collected through a series of textboxes. The Mathematical Model dimension had five textboxes, the Re-Usability & Modifiability dimension had two textboxes, and the Share-Ability dimension had one textbox to complete. The explanations of required focus for the peer feedback within the three dimensions follow.

For the Mathematical Model dimension, the students were required to write feedback concerning the degree to which the teams' math model addressed the complexities imbedded in the problem and the rationales the team used to justify the mathematics used in model. For the *JIT Manufacturing* MEA, the comments with regards to the math model were to address whether or not the evaluated team employed appropriate statistical measures and/or other mathematical processes to yield an inarguable ranking. To assess this aspect, the students are prompted to summarize the mathematical model in their own words, apply the team's procedure and generate results, and give recommendations for improvement. The comments on the rationales were to address whether or not the evaluated team provided sufficient reasoning as to why they chose their mathematical processes. This aspect requires the filling out of two prompts that ask for a summary of rationales provided in the student team solution and recommendations for improvement.

For the Re-Usability & Modifiability dimension, the students were to address the assumptions that the teams made when developing their solution. The assumptions were to be well-communicated in order for the user to understand scenarios under which the procedure could potentially be used in the future and infer possible changes that would enable wider or alternative uses for the procedure. To address this, the peer is challenged to give a summary of the assumptions presented by the team and supply recommendations for improvement.

For the Share-Ability dimension, the students were to address readability issues, such as grammar, missing steps, or poor explanations that make the application of the procedure excessively challenging.

The peers' scores on Mathematical Model: Complexity (scored 1 – 4), Re-Usability & Modifiability (scored 2 – 4), and Audience (Share-Ability): Readability (scored 2 – 4) are the primary source of quantitative data for this analysis (all shown highlighted in APPENDIX A). Peer responses to the four following peer review prompts (shown below as they are displayed on the website for each textbox) are the primary source of qualitative data analyzed in this study. (These prompts within the peer feedback assignment are shown highlighted in APPENDIX B.)

1) Mathematical Model (Complexities): "Provide constructive recommendations on how to better address the complexity of the problem or eliminate errors."
2) Mathematical Model (Rationales): "Provide constructive recommendations to improve the rationales the client needs to understand the procedure."
3) Re-Usability & Modifiability: "Provide constructive comments on anything you feel is missing that would help the client better understand the circumstances under which this procedure can be used."
4) Share-Ability: "Provide constructive recommendations on how to make the procedure easier for the client to use and replicate."

C. Data Analysis

The peer feedback data collected for the *JIT Manufacturing* MEA was analyzed using a mixed-methods approach. The initial quantitative analysis of all 1168 students in the course was used to determine student participation in the peer feedback assignment. The initial analysis compared how many students were in the course with how many students participated in Peer Calibration and how many students completed Peer Feedback on the Second Draft. Next, a mixed-methods analysis was conducted on the completed peer feedback. This analysis includes a quantitative assessment of the peer scores (using the 1 to 4 scale shown in APPENDIX A) and a qualitative analysis of the written peer feedback.

Out of the 1168 students enrolled in the course (~350 teams), 60 randomly selected teams' Second Draft solutions for the *JIT Manufacturing* MEA were scored by an expert grader on all three *MEA Rubric* dimensions. These 60 teams were selected for a previous study in which the teams' work was analyzed by an expert to determine the progression of teams' solutions.[12] This study showed a lack of progress in teams' work after peer feedback. So the same 60 teams work was selected to determine the type of feedback that they received from peers. The expert's scores were based on the *Instructors' MEA Assessment/Evaluation Package* (*I-MAP*) (APPENDIX C), a grading guide seen only by the graders and not by the students during the peer feedback process. Three of these scores were used for this study: Mathematical Model (Complexity) (Table 1), Re-Usability & Modifiability (Table 2), and Share-Ability (Ease of Use) (Table 3). These three scores were quantitatively compared to the equivalent peers' scores to understand the quantitative accuracy of peer feedback.

Table 1. Mathematical Model (Complexities) Score Summary

Score	Description of Requirements
4	Mathematics of Score 3 with clear mathematical rationales and no mathematical errors
3	Accounts for variability and distribution and there is a clear tie breaker strategy. (ex. more than mean & standard deviation)
2	Accounts for variability, but not distribution. (ex. only use mean and standard deviation)
1	Does not account for variability or distribution. (ex. only use mean)

Table 2. Re-Usability & Modifiability Score Summary

Score	Description of Requirements
4	One thing needs work
3	Few things need work (criteria for success, constraints, assumptions, or limitations)
2	Missing all or most of the standard introduction parts of the memo.

Table 3. Share-Ability (Ease of Use) Score Summary

Score	Description of Requirements
4	Easy-to-read-and-use procedure in memo format.
3	Readable and usable, but not in memo format.
2	Procedure is difficult to read and use

The 60 teams received feedback from 152 students; each team received between 1 and 4 peer review(s). The 152 peer reviews were then qualitatively analyzed to understand the nature of the written peer feedback. Analysis of the four written comments was done to elicit themes in the feedback to understand the focus and content of the written feedback for each dimension. It is assumed that this sample group is representative of the population of their fellow classmates.

For the qualitative analysis, the peer feedback was segmented and coded to determine the dimensions to which the feedback applied. Verbal analysis[16] was used to quantify the codes. The coding scheme (Table 4) was based on the *MEA Rubric* dimensions (APPENDIX C) Additional codes were used to identify non-constructive feedback (i.e. "no problems" or "I don't know."), blank feedback (i.e. "---" or "none"), and praise for the current solution (i.e. "good" or "You did this quite well.").

Table 4. *MEA Rubric* Dimensions for Qualitative Analysis

Dimension	Description
Mathematical Model (Complexities)	Feedback that focuses on the statistics utilized or lack of, the tie breaking strategy or lack of, and/or any comments on faults in the current mathematical model.
Mathematical Model (Rationales)	Feedback that focuses on the lack of reasoning behind the mathematics chosen for the procedure.
Re-Usability & Modifiability	Feedback that focuses on criteria for success, constrains, assumptions, and/or limitation. Also feedback that focuses on possible applications of this solution or modifications that would be needed for alternative applications.
Share-Ability	Feedback that focuses on results, sample equations, grammatical errors, content that needs to be further explained or minimized, formatting and/or any comments that are targeted at making the solution more user-friendly.

Sub-codes were developed for feedback coded as Mathematical Model (Rationales) and Re-Usability & Modifiability. For the Mathematical Model (Rationales) sub-dimension two sub-codes were applied: vague and focused feedback. Vague feedback has appropriate content for this dimension, but is not specific to the evaluated team's solution. Focused feedback consists of comments that discuss specific components of the team's procedure that are lacking rationales. For the Re-Usability & Modifiability dimension, feedback sub-codes were "within the scope of the problem" and "not within the scope of the problem". The feedback considered not within the scope of the problem focused on applying the current model to irrelevant scenarios, such as bad weather and poor road conditions. Feedback within the scope of the problem discussed possible effects of applying the procedure to other data relevant to the problem, such as shipping data for longer travel distances or data measured using a different time unit.

IV. Results

A. Student Participation in MEA Peer Feedback Process

The initial analysis of all of the peer feedback data for the class showed that a large group of students did not complete this assignment. Out of the 1168 students, 161 students did not access the assignment and were therefore never assigned to any team for the double-blind peer review. Out of the 1007 remaining students that at least opened the assignment, 173 students did not complete the peer calibration, making them ineligible for the double-blind peer review. Out of the remaining 834 students that completed the peer calibration, 80 students did not give feedback to their assigned peer team. The total number of students that did not participate in the peer feedback process of the *JIT Manufacturing* MEA is 414 out of 1168 (~35%). Out of the 1007 students that were assigned to a team, 253 students (~25.1%) did not give the assigned team feedback. The 60 selected teams were to receive peer feedback from 206 students, but only 152 students gave feedback to these teams. 54 students (~26.2%) did not give the assigned team any feedback.

B. Lack of Constructive Feedback

Out of the 152 students who did give peer feedback, there were many instances where no constructive criticism was provided. This was typically feedback that only gave positive reinforcement or praise. Sometimes this feedback was just simple statements to fill the textbox. The Re-Usability & Modifiability dimension had the most student comments that consisted of no constructive feedback at 44% (67 students). About 25% of the comments associated with each of the other three dimensions did not provide constructive feedback. Specifically, the Mathematical Model (Complexity) dimension had 38 student responses, the Mathematical Model (Rationale) dimension had 36 student responses, and the Share-Ability dimension had 41 student responses that did not consist of constructive feedback. Only 3 students out of the 152 gave no constructive feedback across all 4 dimensions. Most instances without constructive feedback only occurred within 1 or 2 prompts of a student's feedback. Some student responses within the peer feedback that were coded as not providing constructive feedback are shown in Table 5. Even though most students gave some constructive feedback, only 45 students (~29.6%) gave constructive feedback along all four dimensions. Examples of constructive feedback are explained in the discussion of types of feedback (Section D).

Table 5. Peer Feedback without Constructive Feedback (Blank Feedback or Praise)

Dimension	Example of Student Responses without Constructive Feedback
Mathematical Model (Complexity)	"The complexity of the problem is well addressed. I can't think of anything as of now that I could suggest to improve this area." – *Student GQSQ* "Sounds good." – *Student GLJQ* "None." – *Student G2GY*
Mathematical Model (Rationale)	"You did this step quite well." – *Student GSWI* "All of your actions seem rational and are stated appropriately." – *Student HCKC* "Good rationales for the most part." – *Student GQZJ*
Re-Usability & Modifiability	"I'm not sure what could be missing." – *Student GPC2* "Overall good procedure." – *Student GVM3* "Assumptions." – *Student HYBC* "I don't know of anything." – *Student GSXW*
Share-Ability	"Good job team. It is a very easily read piece. Good luck with the final score on your MEA 2." – *Student G4PI* "No problems with share-ability." – *Student GYAO*

C. Peer Scoring vs. Expert Scoring

Figures 2 through 4 show the quantitative comparison of the peer scores versus expert scores. The middle score of P = E refers to the peer score being the same as the expert score. Peer scores that were higher than the expert scores are denoted by P = E + 1, E + 2, or E + 3. Peer scores that were lower than the expert scores are denoted by P = E – 1, E – 2, or E – 3.

The quantitative analysis shows that the students were more lenient than the expert across all three dimensions (Figure 2, Figure 3, and Figure 4). The Mathematical Model dimension had the greatest positive skew in peer feedback scores (Figure 2) – with 97 students (~63.8%) scoring the team solutions higher than the expert. The Re-Usability & Modifiability dimension (Figure 3) was only slightly more positive than the Share-Ability dimension (Figure 4). Seventy students (~46.1%) scored the team solutions higher than the expert on the Re-Usability & Modifiability dimension. Sixty-four students (~42.1%) scored team solutions higher than the expert on the Share-Ability dimension.

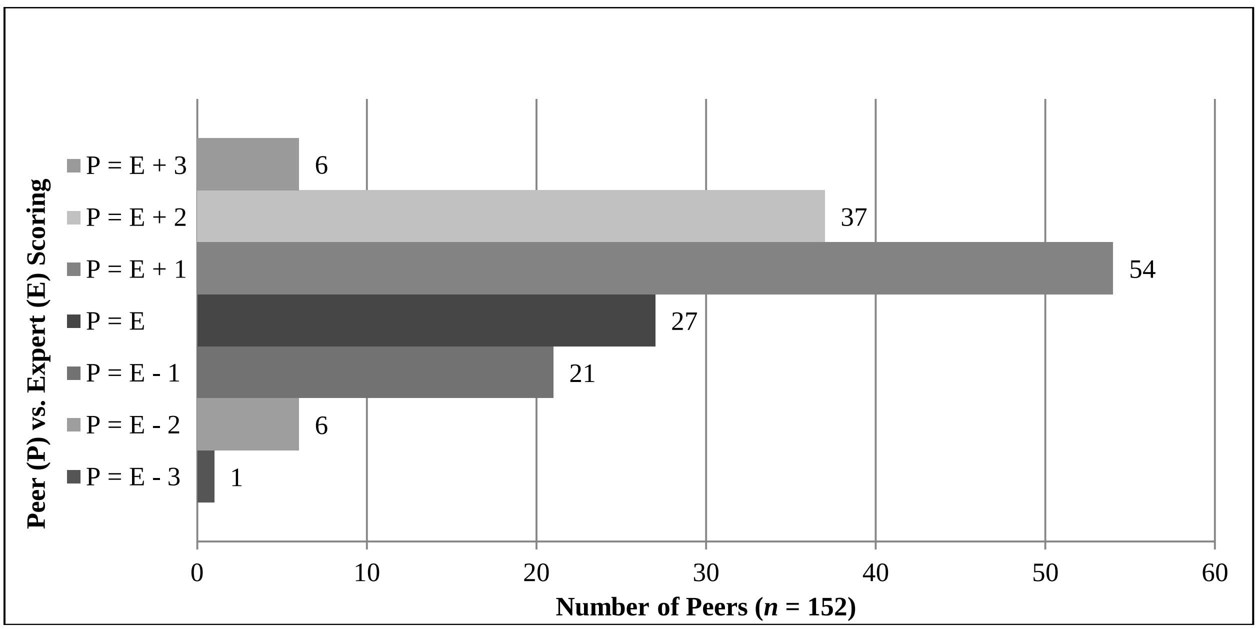

Figure 2: Peer vs. Expert Scores on Mathematical Model (Complexity) Dimension

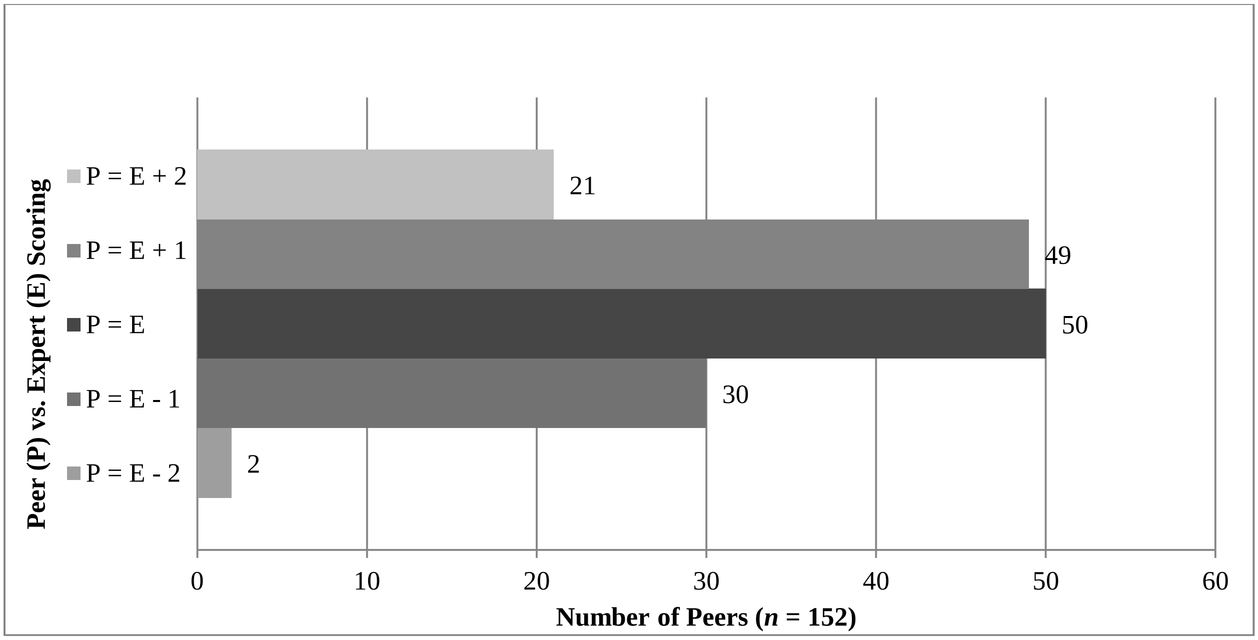

Figure 3: Peer vs. Expert Scores on Re-Usability & Modifiability Dimension

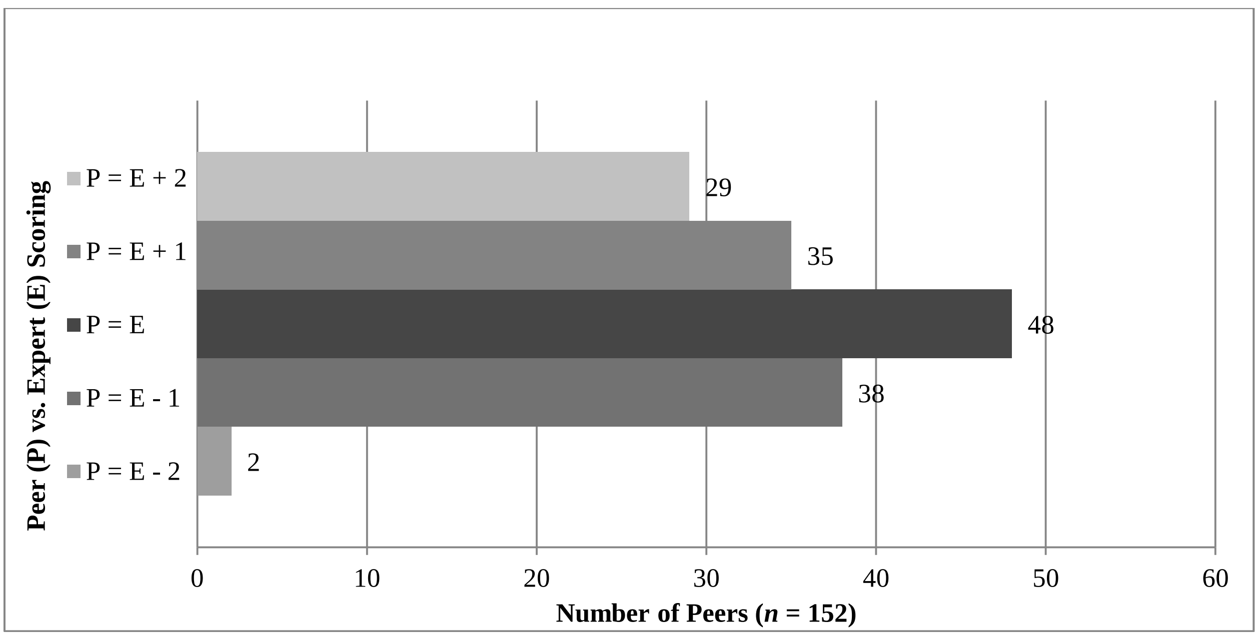

Figure 4: Peer vs. Expert Scores on Share-Ability (Audience): Readability Dimension

D. Written Peer Feedback on the Three Dimensions of the *MEA Rubric*

The qualitative analysis shows that the students who did give constructive feedback typically did not always focus on the appropriate dimension called for by each prompt. This was most common within the Mathematical Model (Complexities) and Re-Usability & Modifiability dimensions. Many of the peer reviews also had comments that were vague and could be applied to almost all models. This was most common within the Mathematical Model (Rationale) dimension. Table 6 shows the types of feedback that were given by the students. The black boxes within Table 6 show the constructive feedback that was appropriately focused on the prompted dimension. The majority of the students only gave constructive feedback in one dimension for each textbox, but almost half of the students that gave constructive feedback also gave praise within the same textbox.

Table 6. Number of Coded Peer Feedback Segments Associated with Each *MEA Rubric* Dimension

		Dimension Feedback was Given Within				
		Mathematical Model (Complexities)	Mathematical Model (Rationale)	Re-Usability & Modifiability	Share-Ability	Total Comments for Each Dimension in all Feedback
Analysis Coding that Feedback was Related to	Praise	47	56	34	62	**199**
	Mathematical Model (Complexities)	**68**	12	25	15	**120**
	Mathematical Model (Rationale)	4	**78**	8	7	**97**
	Re-Usability & Modifiability	3	2	**40**	11	**56**
	Share-Ability	50	18	24	**91**	**183**
	Total Constructive Feedback within Each Dimension	**125**	**110**	**97**	**124**	
	Total Comments within Each Dimension	**172**	**166**	**131**	**186**	

Over half of the students that gave constructive feedback (68 out of 125) on the Mathematical Model (Complexities) dimension gave feedback that was appropriately related to math advice. The majority of the other students who gave constructive feedback on this dimension (50 out of 125) focused on the Share-Ability dimension. The common types of appropriate feedback consisted of advice to use more statistical measures and/or add a tie breaker to the procedure. The common types of constructive feedback that were inappropriately located were advice to add a sample calculation to encourage greater user understanding of the procedure and/or to reword or reorganize the procedure for greater ease of use. Some samples of students' appropriate and ill-located peer feedback within the Mathematical Model (Complexities) dimension are shown in Table 7.

Table 7. Peer Feedback on the Mathematical Model (Complexities) Dimension

Type of Peer Feedback	Examples of Peer Feedback
Focused only on Mathematical Model (Complexities)	"There are many other things like standard deviation and median that can help you better support your conclusion." – *Student HVDP* "Give a better tie breaker. DDT cannot pick more than one company so your tie breaker does not sufficiently break a tie." – *Student GPUG* [This team's tie breaker was a written statement for DDT to 'just pick one'.]
Share-Ability Feedback	"A little confusing. Maybe more wording. I got the right answer with different numbers. Be more specific. Instead of saying 'multiply the result from step 5' actually put what that is to avoid confusion." – *Student G2RX* "For the client to better understand the procedure, a sample calculation would be nice." – *Student HM6B*

Seventy-eight students of the 110 who gave constructive feedback within the Mathematical Model (Rationale) sub-dimension appropriately focused on the team's mathematical justifications, although they were commonly vague suggestions that restated what a rationale should be and were not tailored to the reviewed team solution. Table 8 shows some examples of students' reviews that are focused to the team's solution and others that are vague recommendations.

Table 8. Examples of Peer Feedback on the Mathematical Model (Rationale) Dimension

Level of Detail	Examples of Peer Feedback
Focused (Specific to Team's Solution)	"The team could use rationales to explain why they used only the average and also why they decided on their tie breaker." – *Student GOUD* "Explain further your rationales for using standard deviation and mean." – *Student GZSD* "You have not specifies why you have created the bins in the way you have like why 0 – 7, 8 – 15, and so on. You also have not given rationales as to why the accuracy is more important than precision." – *Student GQ3Y*
Vague (Generic)	"Explain explicitly why you did things the way you did." – *Student HJ53* "Supply more rationales throughout. Each decision should be supported by rationale as to eliminate any question of your method." – *Student GRJP*

The Re-Usability & Modifiability dimension had the greatest variation of responses that were either inappropriate applications within the appropriate dimension, focused on Share-Ability dimension aspects (~25% of students who gave constructive feedback), or focused on Mathematical Model (Complexities) dimension components (~25% of students who gave constructive feedback). Sample feedback is given in Table 9.

Table 9. Examples of Peer Feedback on the Re-Usability & Modifiability Dimension

Dimension Related To	**Examples of Peer Feedback**
Re-Usability & Modifiability (Relevant Applications)	"How is the client supposed to know how to set a bound/limit for a different data set with much larger or smaller values?" – *Student GLYT* "An assumption that the data is calculate in minutes could be a good addition." – *Student GN4M* "Do you explain that this procedure will work if the companies are shipping for 50, 100, or 800 miles away?" – *Student GKHD*
Re-Usability & Modifiability (Applying Model to Scenarios Outside of the Scope of the Problem)	"One assumption that could be added is that all of the companies use the same type of delivery vehicles. It might not necessarily be that important, but it could come in to play in some circumstances." – *Student GQSQ* "You may also want to assume that the client has some software to calculate the standard deviation of a large dataset." – *Student GQ3Y* "While creating the third draft, I would consider the assumption that the outlier are very late arrival times due to weather, a flat tire, or something of that sort, because that might be the case most of the time, but not all of the time." – *Student G5BO*
Share-Ability	"Make it less complicated and time consuming." – *Student GNBJ* "Put your letter in memo form. This will show more organization." – *Student GS2M*
Mathematical Model (Complexities)	"Should you use outliers? What information will you use or not use? How will you determine a winner if two companies get the same point value? Is there a way to ensure that the point value system will always have a winner and no ties?" – *Student G3EY* "I feel that perhaps the mean should be considered and is missing from your procedure. Also maybe standard deviation could be utilized to determine outliers. A tie-breaker is necessary as well." – *Student GQD5*
Mathematical Model (Rationales)	"They could add why they chose to do what they did." – *Student GOUD* "In the event of a tie, why are they equally good choices?" – *Student G2H3* "The only thing is the extreme outlier procedure is not justified as to why this solution and not another." – *Student GS24*

Ninety-one students gave constructive feedback appropriately related to the Share-Ability dimension, but some students also used this space to summarize all of their previous feedback. Examples of both of these common types of feedback are shown in Table 10.

Table 10. Examples of Peer Feedback on the Share-Ability Dimension

Focus of Feedback	Examples of Peer Feedback
Share-Ability	"The procedure is workable but you can consider clarifying what will happened in the event of tie. But overall written in a very clear way." – *Student G6J6* "Give calculations along with your results." – *Student GZSD* "Provide example calculations and formulas within your procedure." – *Student GXK3* "Break it down into numbered steps to make it clear and more organized for the users eyes." – *Student GVM3* "Some of the stuff explained in the introduction could be shortened to a few sentences. Overall, however it seems like a pretty good, well-thought-out method." – *Student GN4M*
Summary (including Share-Ability)	"As stated above, the only problem with the procedure was how to break a tie. I would recommend that the team try to make this process more clear in explanation and strategy." – *Student G2S7* "The team can eliminate the point system and only consider adding the results which makes it easier to produce results. This also give more accurate results." – *Student G4Y5* "Be less wordy in description. Provide assumptions about the data points. Otherwise very clear and able to be used on any data points." – *Student G7ZC* "Use more justifications. Add some more rationales. Clarify and elaborate!" – *Student G2RX*

V. Discussion

The peer feedback process for MEAs is worth a small fraction of the students' final grades and relies on "Good Faith Effort". [13] Supervision and in-depth grader review of all peer feedback is just too costly for such a large class size. Still it may be useful to consider strategies for increasing engagement in the peer review process. An obvious strategy would be greater grade incentives for participation. However, such a practice may not increase the amount or quality of constructive feedback. Another strategy may be to ask the student teams to rate the value of the feedback they receive. This actually has been tried and resulted in participation problems and scoring reliability issues.[11] Still another strategy might be to consider linking grade incentives to whether or not the evaluated team actually improves their model based on their peer feedback. A more learning-focused strategy may be to help students come to value the peer feedback process as part of their preparation for the engineering profession through discussion of this topic and modeling of appropriate feedback strategies.

The quantitative analysis shows that peers have a tendency to score more leniently than the expert on all three evaluated dimensions of the MEA. The higher peer scores may be due to a lack of understanding of each dimension. This seems to bare out in the qualitative data analysis. Students were confused as to what aspects of feedback are specific to the Mathematical Model (Complexities) dimension versus the Share-Ability dimension. Further, for Re-Usability &

Modifiability, the majority of students had comments related to other dimensions or comments that were not within the scope of this problem.

The students' inability to differentiate the Share-Ability and Mathematical Model (Complexities) dimensions could be mitigated by providing instruction on what constitutes Mathematical Model feedback versus general writing feedback. This would enable students to make more progress on the quality of their math models following peer review, which has previously been shown to be limited.[12] When devising this instruction, there could be a more in-depth analysis of the math language used in the *I-MAP* as compared to students' own math wording. This wording may give some insight on ways for peers and instructors to address misconceptions in terms students better understand. For instance, when giving peer feedback on teams' solutions that only used the mean to solve the *JIT Manufacturing* MEA, some students discuss the lack of precision (possibly referring to variability and distribution) and only focusing on accuracy (possibly referring to central tendency). The words precisions and accuracy are not used in *I-MAP* but may be familiar to students from other coursework.

The Re-Usability & Modifiability dimension appears to be the dimension for which students need the greatest explanation on how to help their peers write applicable assumptions. In prior semesters, the scoring of this dimension was also found to be difficult for GTAs, so the dimensions were refined. Previously, this dimension had a lot more overlap with the Share-Ability dimension.[15] The dimensions were redefined to have Share-Ability focus on the ease of use and language instead of modifiability. Modifiability was combined with the generalizability component to create the updated Re-Usability & Modifiability dimension. [15] The updated version of the dimensions appears to have greatly improved the reliability of the Share-Ability dimension; this dimension received the most constructive feedback and had the least positive skew in scores. Still, there appears to be confusion over the purpose of the Re-Usability & Modifiability dimension. Further revision of the Re-Usability & Modifiability dimension that teases these two ideas apart could improve the quality of peer feedback. Also, the provision of examples of what assumptions would be within the scope of the problem and what would be irrelevant to the current problem may help students better understand this dimension.

This analysis also reveals that many students understood the purpose of the Mathematical Model (Rationales) dimension, but did not provide an in-depth review of the team's solution. To encourage students to give feedback tailored to the solution they are reviewing, it may be beneficial to help peers understand the effectiveness of focused feedback over vague feedback.[9]

Finally, it may be advisable to remove the quantitative scoring from the peer feedback process. Then teams will not be misled by high scores. Their focus would be on the feedback they receive. Focusing on qualitative, formative feedback over numeric scoring has been acknowledged by others to be an effective technique in the peer feedback process, including an engineering undergraduate program.[17] This method could be implemented and tested to learn whether teams respond (change/improve their models) more when only qualitative peer feedback is provided.

VI. Conclusion

This study was an initial attempt to understand the nature of peer feedback occurring during the implementation of MEAs, open-ended mathematical modeling problems. It was confirmed that full student participation in peer feedback is problematic. Student scoring of other team's work is more lenient than that of an expert, perhaps due to a lack of understanding of the *MEA Rubric* dimensions. This is somewhat confirmed by students' provision on feedback in one dimension that is really intended for another dimension. Potential strategies for improving student participation in peer feedback were discussed. Means for improving students' understanding of dimensions along which they are to assess peer work and provide peer feedback were also discussed.

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. Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77,* 81-112.
2. Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119-144.
3. Carless, D., Slater, D., Yang, M., & Lam, J. (2010). Developing sustainable feedback practices, *Studies in Higher Education*, *36*(4), 395-407.
4. National Academy of Engineering. (2004). *The engineering of 2020: Visions of engineering in the new century*, Washington, DC: The National Academic Press.
5. Institute of Electrical and Electronics Engineers. IEEE Policies, Section 7 – Professional Activities (Part A – IEEE Policies), 7.8 Code of Ethics. Retrieved from http://www.ieee.org/about/corporate/governance/p7-8.html.
6. McCarthy, A. M., & Garavan, T. N. (2001). 360* feedback processes: performance improvement and employee career development. *Journal of European Industrial Training*, 25, 5-32.
7. Muller, L. (2001). The importance of peer review. *Australian Journal of Dairy Technology, 56*(3), 191-191.
8. ABET Engineering Accreditation Commission. (2010). Criteria for accrediting engineering programs. ABET Inc.: Baltimore, MD, Retrieved from http://www.abet.org/forms.shtml#For_Engineering_Programs_Only.
9. Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science*, *27*(4), 375-401.
10. Diefes-Dux, H. A., Hjalmarson, M. A., Miller, T. K., & 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. 17-35). Rotterdam, the Netherlands: Sense Publishers.

11. Diefes-Dux, H. A. & Verleger, M. A. (2009). Student reflections on peer reviewing solutions to model-eliciting activities. *Paper presented at the 39th ASEE/IEEE Frontiers in Education Conference.* San Antonio, Texas.

12. Carnes, M. T., Cardella M. E., & Diefes-Dux H. A. (2010). Progression of student solutions over the course of a model-eliciting activity (MEA). *Paper presented at the 40th ASEE/IEEE Frontiers in Education Conference,* Washington DC.

13. Diefes-Dux, H. A. & Imbrie, P. K. (2008). Chapter 4: modeling activities in a first-year engineering course. In Zawojewski, J. S., Diefes-Dux, H., & Bowman, K. (Eds.), *Models and modeling in engineering education: designing experiences for all students* (pp. 36-92). Rotterdam, the Netherlands: Sense Publishers.

14. Verleger, M., Diefes-Dux, H., Ohland, M.W., Besterfield-Sacre, M., & Brophy, S. (2010). Challenges to Informed Peer Review Matching Algorithms, *Journal of Engineering Education*. *99*(4), 397-408.

15. Diefes-Dux, H. A., Zawojewski, J. S., & Hjalmarson, M. A. (2010). Using educational research in the design of evaluation tools for open-ended problems. *International Journal of Engineering Education*, *26*(4):807-819.

16. Chi, M. T. H. (1997).Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of Learning Sciences*, *6*(3), 271-315.

17. O'Moore, L. M., & Baldock, T. E. (2007). Peer assessment learning sessions (PALS): an innovative feedback technique for large engineering classes. *European Journal of Engineering Education*, *32*, 43–55.

APPENDIX A: *MEA RUBRIC* WITH THREE DIMENSIONS

Dim.	Item Label	Full Item Wording	Score	
Mathematical Model	No Progress	No progress has been made in developing a model. Nothing has been produced that even resembles a poor mathematical model. For example, simply rewriting the question or writing a "chatty" letter to the client does not constitute turning in a product.	True	0
			False	4
	Mathematical Model Complexity	The procedure fully addresses the complexity of the problem.	4	
		A procedure moderately addresses the complexity of the problem or contains embedded errors.	3	
		A procedure somewhat addresses the complexity of the problem or contains embedded errors.	2	
		Does not achieve the above level.	1	
	Data Usage	The procedure takes into account all types of data provided to generate results OR justifies not using some of the data types provided.	True	4
			False	3
	Rationales	The procedure is supported with rationales for critical steps in the procedure.	True	4
			False	3
Re-Usability & Modifiability	Re-Usability/ Modifiability	The procedure not only works for the data provided but is clearly re-usable and modifiable. Re-usability and modifiability are made clear by well articulated steps and clearly discussed assumptions about the situation and the types of data to which the procedure can be applied.	4	
		The procedure works for the data provided and might be re-usable and modifiable, but it is unclear whether the procedure is re-usable and modifiable because assumptions about the situation and/or the types of data that the procedure can be applied to are not clear or not provided.	3	
		Does not achieve the above level.	2	
Audience (Share-Ability)	Results	Results from applying the procedure to the data provided are presented in the form requested.	True	4
			False	1
	Audience Readability	The procedure is easy for the client to understand and replicate. All steps in the procedure are clearly and completely articulated.	4	
		The procedure is relatively easy for the client to understand and replicate. One or more of the following are needed to improve the procedure: (1) two or more steps must be written more clearly and/or (2) additional description, example calculations using the data provided, or intermediate results from the data provided are needed to clarify the steps.	3	
		Does not achieve the above level.	2	
	Extraneous Information	There is no extraneous information in the response.	True	4
			False	3

Overall Assessment is the lowest score on any rubric item.

APPENDIX B: EXCERPTS FROM THE PEER FEEDBACK ASSIGNMENT

Dimension	Peer Feedback Instructions (Excerpt of Assignment)
Mathematical Model	Excerpt of Instructions Students Received for Peer Feedback on Mathematical Model: a. Under the Mathematical Model heading in the textbox provided, briefly summarize in your own words the mathematics used in the procedure. b. To provide a critical review, you must actually try to apply the team's procedure. Use the historical data stored as **justintime_data_hmwk9.txt** on Blackboard. There are eight shipping companies to consider. Be sure to use the procedure as written. In the textbox provided, state the results found by applying the procedure as written to the historical data for the ***eight*** shipping companies, providing some quantitative results. Also, describe any problem(s) you experienced when trying to apply the procedure. c. In the third textbox in the Mathematical Model section, provide constructive recommendations on how to better address the complexity of the problem or eliminate errors. d. With regards to the rationales, check whether the team provided rationales for the critical steps in the procedure (Do you completely understand why the team believes the critical steps are correct and/or necessary?). Briefly summarize the rationales provided in the team's memo in the textbox provided. e. Then, in the next textbox, make recommendations to improve the rationales so that the client can better understand why the procedure is correct.
Re-Usability & Modifiability	Excerpt of Instructions Students Received for Peer Feedback on Re-Usability/Modifiability: You will also critically evaluate the **Re-Usability and Modifiability** of the team's work. Select the level to which you feel the team's procedure is re-usable and modifiable. Recall these definitions: • *Re-usability* means that the procedure can be used by the client for new but similar situations • *Modifiability* means that the procedure can be modified easily by the client for slightly different situations Well stated assumptions and limitations of the procedure are needed to achieve re-usability and modifiability. a. In the first textbox, summarize the assumptions the team made. b. In the next textbox, provide constructive comments to the team on how to better specify the conditions under which the procedure can be applied.
Audience (Share-Ability)	Excerpt of Instructions Students Received for Peer Feedback on Share-Ability: Finally, you will critically evaluate how well the team attended to the needs of the **Audience (Share-ability)**. Recall this definition: • *Share-ability* means that the procedure can be used by the client to reproduce results. a. In the textbox give feedback on whether the team presented the results (remember that both team winners and quantitative results are expected) and/or included any extraneous information in the memo. Then support your comments by making recommendations to make the procedure easier for the client to use and replicate.

APPENDIX C: I-MAP (JUST-IN-TIME MANUFACTURING MEA FEEDBACK AND ASSESSMENT – CORE ELEMENTS OF PERFORMANCE ON AN MEA)

Mathematical Model ***(Specific to Just-In-Time Manufacturing MEA)***

Looking beyond a single measure of central tendency: This particular MEA is set in a context where patterns of late arrival are important. Therefore, the data sets are designed so that the differences in the mean are insignificant. This is intended to nudge students to look beyond measures of central tendency. Therefore, more than one statistical measure is needed. Teams might use a number of measures simultaneously, or one following the other. They might also use one measure to produce an answer and another to "check" how well the answer works, leading to a possible revision. Results from statistical procedures may be aggregated in some fashion using rankings, formulas, or other methods.

In a high quality model:

- *The procedure looks past measures of central tendency and variation to look at the actual distribution of the data, where attention is drawn to the frequency of values, particularly minimum and maximum values.*
- *Final overall ranking measure or method must be clearly defined. Completes the sentence, the ranking procedure is based on…*
  - *This is Part B of the standard introduction :*

    B. Describe what the procedure below is designed to do or find – be specific (~1- 2 sentences)
- *Critical steps that needs justification / rationale:*
  - *When teams use any statistical measures, these measures must be justified – explain what these measures tells the user.*
  - *When developing intermediate ranking or weighting methods, these must be justified.*

LEVEL 1 -

- The procedure described does **not account** for both the variability or distribution of these data. Students cannot move past this level if only the mean of the data is used in their procedure.
- Merely computing a series of statistical measures without a coherent procedure to use the results fall into this level.

LEVEL 2 –

- The procedure described accounts for the variability, but not the distribution, of these data.
- Mathematical detail may be lacking or missing.
- Mathematical errors might be present.
- If the solution demonstrates lack of understanding of the context of the problem, this is the highest level achievable.
- If there is an indication that the team does not understand one or more statistical measures being used, drop to the next level

LEVEL 3 –

- The procedure described accounts for both the variability and distribution of these data. That is the procedure includes more than the mean and/or standard deviation. The ranking procedure accounts for how the data is distributed.
- The procedure provides a viable strategy for how to break tie.
- Some mathematical detail may be lacking or missing.

- Mathematical errors might be present.
- If there is an indication that the team does not understand one or more statistical measures being used, drop to the next level

LEVEL 4 –

- Clear statement of what defines the overall ranking.
- Mathematical detail should be clear from start to finish.
- Mathematical errors should be eliminated.

Additional but separate LEVEL 4 criteria:

- Rationales for the critical steps in the procedure must be provided. (If the rationales provided are not correct, this is FALSE. If they just need minor clean-up/clarification this is TRUE.)
- If all data provided is not used in the mathematical model, this must be explained or justified. (If the justifications are not correct this is FALSE. If they just need minor clean-up/clarification this is TRUE.)

Re-Usability and Modifiability

The mathematical model produced must be *Re-usable* (the client can use it for new but similar situations) and *Modifiability* (the client can modify it easily for slightly different situations). Generally, one would not produce a mathematical model to solve a problem for a single situation. A mathematical model is produced when a situation will arise repeatedly, with different data sets. Therefore, the model needs to be able to work for a variety of data sets. The model may be in the form of a procedure or explanation that accomplishes a task, makes a decision, or fills a need for a client.

Further, a useful mathematical model is adaptable to similar, but slightly different, situations. For example, a novel data set may emerge that wasn't accounted for in the original model, and thus the user would need to revise the model to accommodate the new situation. Thus, one should strive for clarity, efficiency and simplicity in mathematical models; as such models are the ones that are more readily modified for new situations.

At a minimum, the mathematical model should include assumptions about the situation and the types of data to which the procedure can be applied. Hard-coded quantitative values imbedded in a procedure require explicit assumptions or explanations.

If the mathematical model is not developed in enough detail to clearly demonstrate that it works on the data provided, it cannot be considered re-usable and modifiable.

Student teams should state that the procedure is designed to rank shipping companies in order of best to least able to meet DDT's timing needs given historical data for multiple shipping companies of time late for shipping runs between two specified locations.

Students should also indicate limitations of their procedure. Limitations might be centered around hard-coded quantitative values imbedded in a procedure. These require explicit assumptions or explanations. Hard-coded values might include: an indication of what is considered late, ways to parse the data (related to degree of lateness), weighting factors.

Level 2 – Missing all or most of the standard Introduction parts A & C. (Part B is part of the *mathematical model criteria)*

I. Introduction

A. In your own words, restate the task that was assigned to your team (~1-2 sentences). This is your team's consensus on who the client is and what solution the client needs.

C. State your assumptions about the conditions under which it is appropriate to use the procedure. Another way to think about this is to describe the limitations of your procedure.

Level 3 – ~2-3 things need work, typically from the standard introduction (criteria for success, constraints, assumptions, or limitations) or implicit assumptions
Level 4 - ~1 thing needs work.

Audience (Share-ability)

***Effectively communicating to the client*: The mathematical model is share-able – the client can use it to reproduce results.**

Although the client (or an intermediary) has "hired" the consultant team to construct a mathematical model, the client (or the intermediary) needs and wants to understand what the model accomplishes, what trade-offs were involved in creating the model, and how the model works. A high quality product (i.e., model communicated to the client) will clearly, efficiently and completely articulate the steps of the procedure. A high quality product will also illustrate how the model is used on the given set of data. The description will be clear and easy to follow; it must enable the results of the test case to be reproduced. Given this type of information, the client will be able to intelligently use and/or modify the model for new situations. At a minimum, the results from applying the procedure to the data provided must be presented in the form requested.

RESULTS: Results of applying the procedure MUST be included in the memo. This must include a ranking of all shipping companies (or listing of those discarded prior to ranking) and quantitative (possibly intermediate) results. If results are missing students will receive a Level 1 (D grade) for the MEA.

PROCEDURE: The client requires a relatively easy-to-read-and-use procedure. If this has not been delivered, the solution is not Level 3 work.

If you, as a representative of the client, cannot replicate or generate results, the solution is not Level 3 work.

Memos left in outline form may only receive a maximum Level 3 audience rating.

EXTRANEOUS INFORMATION might include mentions of specific tools (MATLAB or Excel) to complete computations or overly describing how to compute basic statistical measures (e.g. mean, standard deviation).