

Marquette University
e-Publications@Marquette

Mathematics, Statistics and Computer Science
Faculty Research and Publications

Mathematics, Statistics and Computer Science,
Department of

1-1-2015

Characteristics of Feedback that Influence Student Confidence and Performance during Mathematical Modeling

Hyunyi Jung

Marquette University, hyunyi.jung@marquette.edu

Heidi A. Diefes-Dux

Purdue University

Aladar K. Horvath

Ivy Tech Community College - Muncie

Kelsey Joy Rodgers

Purdue University

Monica E. Cardella

Purdue University

Published version. *International Journal of Engineering Education*, Vol. 31, No. 1A (2015): 42-57.

[Permalink](#). © 2015 Tempus Publications. Used with permission.

Hyunyi Jung was affiliated with Purdue University at the time of publication.

Characteristics of Feedback that Influence Student Confidence and Performance during Mathematical Modeling*

HYUNYI JUNG

Mathematics Education, College of Education, Purdue University, Beering Hall of Liberal Arts and Education, R4135, 100 N. University St. West Lafayette, IN 47907-2098, USA. E-mail: jung91@purdue.edu

HEIDI A. DIEFES-DUX

School of Engineering Education, Purdue University, West Lafayette, IN 47907-2045, USA. E-mail: hdiefes@purdue.edu

ALADAR K. HORVATH

General Studies, Ivy Tech Community College, Muncie, IN, 47302, USA. E-mail: ahorvath5@ivytech.edu

KELSEY JOY RODGERS, MONICA E. CARDELLA

School of Engineering Education, Purdue University, West Lafayette, IN 47907-2045, USA.
E-mails: krodger@purdue.edu, mcardell@purdue.edu

This study focuses on characteristics of written feedback that influence students' performance and confidence in addressing the mathematical complexity embedded in a Model-Eliciting Activity (MEA). MEAs are authentic mathematical modeling problems that facilitate students' iterative development of solutions in a realistic context. We analyzed 132 first-year engineering students' confidence levels and mathematical model scores on a MEA (pre and post feedback), along with teaching assistant feedback given to the students. The findings show several examples of affective and cognitive feedback that students reported that they used to revise their models. Students' performance and confidence in developing mathematical models can be increased when they are in an environment where they iteratively develop models based on effective feedback.

Keywords: confidence; feedback; mathematical model; performance

1. Introduction

Feedback is generally regarded as an important factor in improving students' knowledge and skills [1–3]. When students are given feedback and encouraged to think in different ways, they are able to change their thinking in terms of inquiry, exploration, and creativity [4]. As feedback is used as a method to communicate with learners about how to modify their work, effective feedback positively influences student achievement and motivates their learning [5, 6]. However, several studies related to feedback and its relation to student learning and performance show no consistent pattern of findings over the past 50 years [7]. Conflicting findings might be the result of different settings and methods used in the research. For instance, the nature of feedback given on students' multiple-choice solutions would be different from that of open-ended problems. Results from qualitative analysis of feedback could be different from those from quantitative or mixed-methods analysis. In this study, we qualitatively analyzed teaching assistant (TA) feedback given to first-year engineering students who worked on a Model-Eliciting Activity (MEA).

MEAs are real-world, mathematical modeling problems that require teams of students to develop their solutions through an iterative process [8, 9]. Students can develop higher-order thinking and problem solving abilities by collaboratively solving real-world problems [10]. Students can also build their knowledge when their instructors scaffold their thinking and encourage communication among their peers [11]. When engaging in MEAs, students create their solutions, receive feedback from others, and self-assess their initial solutions to develop an improved product [8, 12]. In the process of revising their solutions, feedback has a crucial role in scaffolding students' critical thinking about the mathematics used. Engagement in the processes of solving and revising problems in teams improves students' collaboration, communication, problem solving, and critical thinking skills, which are called for by the National Council of Teachers of Mathematics [13], the Common Core State Standards for Mathematics [14], and ABET [15].

As students improve their problem solving and critical thinking abilities, they might increase their levels of confidence in addressing the complexity inherent to a problem. Confidence is generally

considered crucial to enhancing students' motivation to achieve their goals and promoting positive attitudes toward learning [16, 17]. In a study of students' confidence on solving problems, Kulhavy [18] found relationships between students' initial confidence levels and their willingness to use feedback. While this study was conducted on an activity where participants solved multiple-choice questions and received feedback on whether their answers were correct [18], we examined feedback on several aspects of students' mathematical model work. Specifically, our research questions were the following: (a) What characteristics of feedback influenced students' confidence and performance in addressing the complexity in a mathematical modeling problem? and (b) How did different types of feedback impact students' confidence and model development?

2. Literature review

This section includes a review of educational studies that address the origin and features of MEAs. We also discuss existing studies on the effects of feedback on students' performance and confidence. The review of the research studies is connected to the background and the rationale for conducting this study.

2.1 Model-eliciting activities (MEA)

MEAs were originally developed by faculty in mathematics education [9] and have been modified and used in other programs such as first-year engineering [19]. MEAs facilitate students' learning to iteratively develop solutions in real-world contexts as they repeatedly improve their solutions through self-assessment and feedback to create a final written document to share with others [8, 12]. As compared to most conventional problem solving activities that require students to engage in choosing a strategy (e.g., operation) to solve a problem with well-specified givens and goals [20], MEAs require students to interpret both the givens and goals from within an ambiguous or undefined problem [21]. When presented with a problem that lacks enough specificity, students need to diagnose the given situation and acknowledge conflicts and alternative interpretations; this process leads students to continuously evaluate and modify their models [20].

The MEA used in this study, *Just-in-Time Manufacturing (JITM) MEA*, was developed to encourage students to apply their mathematical and statistical knowledge to solve an open-ended problem [22, 23]. It has been revised multiple times for use in a first-year engineering course. Data derived from student work and instructional team scoring

and feedback have been used to explore various facets of problem solving and teaching with open-ended problems. The *JITM MEA* requires teams of three or four students to collaboratively develop a mathematical model that employs descriptive statistics to make a procedure to rank shipping companies given historical data. Lengthier descriptions of the *JITM MEA* and its *Instructor's MEA Assessment/Evaluation Package (I-MAP)*, which is a guide for assessing student work, are provided in past publications [24, 25]. The initial client memo to the student teams and the portion of the *I-MAP* that was used to evaluate students' mathematical models are provided in Appendices A and B, respectively.

In a study of students' draft work on the *JITM MEA* after peer feedback (Draft 1 to 2) and after TA feedback (Draft 2 to Team Final Response), Carnes, Cardella, and Diefes-Dux [25] found that more revisions were made after students received TA feedback than after peer feedback. One of the reasons may be that TAs address in their feedback all aspects of a high quality solution, including students' use of statistical methods (e.g., central tendency, variation, and distribution), as encouraged by the *JITM I-MAP*. Even though students made more changes after Draft 2, half of the teams made only minor changes after receiving TA feedback. For example, by the Team Final Response, most teams only used mean and standard deviation as their data-related tools even though the complexity of the problem necessitates consideration of the distribution of the data.

Based on this result, Carnes et al. [25] recommended further investigation into the nature of feedback provided by the TAs. We focused on the students' limited use of feedback for the revision of their mathematical models, as well as the cases where students fully applied TA feedback. We aimed to identify which aspects of TA feedback students used and suggest how to improve TA feedback.

2.2 Feedback and self-confidence

In this section, we conceptualize feedback and self-confidence, and discuss the relationships between feedback and students' self-confidence. Shute [7] defined feedback as information provided to learners that is aimed at changing their views or ideas. Hattie and Timperley [26] defined feedback as knowledge informed by "an agent" such as other people, concerning features of one's practice. Slavin [27], as well as Palincsar and Brown [28], described feedback based on Vygotsky's [29] developmental perspective. Vygotsky proposed that children develop problem solving skills through the guidance of capable peers after first experiencing the problem. This happens within children's zones of proximal

development, the distance between one's independent problem solving ability and the potential level of problem solving ability under the guidance of capable people [29].

As students improve their problem solving ability, they might increase their level of confidence. Self-confidence is the self-evaluation of proficiency [17] and is generally regarded as a valuable asset for students because it enhances their motivation to pursue their goals [16]. Self-confidence also positively influences one's learning achievement and attitudes toward learning; therefore, a good classroom environment promotes students' self-confidence [17]. While American schools emphasize the importance of helping students to increase their self-confidence, they are often questioned on whether students' confidence is overly high considering their actual knowledge [16]. Students' overconfidence might influence the way they participate in an activity that requires several revisions.

In a study of students' confidence on multiple-choice questions, Kulhavy [18] claimed that the interaction between students' initial confidence levels and the correctness of their answers affected students' perceptions when they engaged in the next related activity. When students' confidence levels were high and the correct answer was chosen, students did not pay much attention to the details of the feedback they were given. Melching [30] also supported this idea, stating that students were less likely to request feedback when they were more confident about their answers. On the other hand, when students' confidence was high but their answer was wrong, students were likely to pay special attention to the feedback in order to find their mistakes [18]. Further, when students' confidence in their response was low, it was assumed that students might not be able to understand the material, the problem, or both. Therefore, students were less likely to have the right answers for the activity. In this case, they might need strategies to address their initial struggles. Consequently, feedback had little effect on their revisions regardless of the accuracy of their answers [18].

Such results from Kulhavy's study [18] were based on an activity where participants answered multiple-choice questions and rated their confidence level on a five-point scale without explanation of why they selected both responses. Feedback received by the participants informed them whether their answers were correct. This current study goes beyond simply validating the correctness of students' responses. We believe that students improve their work when they consider the feedback to be helpful and take action as a result of it. Since students in this study received TA written feedback provided via a web based system, there was no face-

to-face interaction while students were given feedback. The reflective nature of written feedback and the nature of online feedback might influence students' model development and confidence levels; therefore, it is crucial to look at students' explanations for their confidence levels to determine whether students attributed their confidence to the feedback they used to revise their mathematical models.

3. Theoretical framework

Characteristics of feedback have been described and discussed by several educational researchers [31, 32]. Nelson and Schunn [33] introduced two characteristics of written feedback: affective and cognitive. Affective feedback includes the tone of the feedback, such as positive, negative, or neutral. Cognitive feedback addresses the processes of knowing and perceiving, such as those involved in solving problems. We also examined other characteristics of feedback, such as redundancy and reproduction to identify feedback that is repetitive or copied directly from another source (e.g., the *I-MAP*), respectively. Table 1 summarizes each feedback characteristic with specific examples.

3.1 Affective feedback

Affective feedback consists of emotional language, such as praise or personal judgment. The type of affective language used in feedback may influence students' responses in positive or negative ways [33]. Since students receive feedback online during a MEA, how students interpret affective aspects of TA feedback might be different from what the TA intends. In this study, TA feedback was considered "neutral" unless it included praise or obviously negative feelings. Similarly, Nelson and Schunn [33] define neutral feedback as language and matter-of-fact statements used to characterize a problem or solution. Positive feedback refers to affective language including approval of work, such as praise. Praise has been regarded as effective in some research [34] but not in other research [32]. Negative feedback has been defined as unproductive criticism with the focus being more personal than task oriented. Kluger and DeNisi [35] suggest that feedback address specific features of learners' work, not the learners themselves.

3.2 General cognitive feedback

General cognitive feedback includes six characteristics: summary, problem, solution, explanation, clarification, and reflection. Summary, problem, solution, and explanation were adopted from Nelson and Schunn [33]'s study, whereas clarification and reflection were adopted from Driscoll [36]'s

Table 1. Type and characteristics of feedback

Type	Characteristics	Description	Sample feedback segment from the <i>JITM MEA</i> data
Affective	Positive	Complimentary statements showing approval of the work or praise.	Great job of addressing both variability and distribution.
	Negative	Negatively phrased criticism without constructive suggestions.	Your procedure is so confusing and I cannot follow it.
	Neutral	Descriptive statements with no obvious emotional aspect, but a matter-of-fact.	You have taken care of the distribution aspect of the data.
Cognitive General	Summary	Addressing claims or statements.	Data is observed and the number of hours late is calculated for each company.
	Problem	Identifying something that are missed or needed to fix.	You have not taken care of the variability of the data.
	Solution	Suggesting possible ways to improve the work.	You need to account for a tiebreaker situation if it arises.
	Explanation	Providing reasons or further description regarding the problem or solution.	That is the frequency of data within certain intervals.
	Clarification	Asking questions to clarify what the students mention.	What are these groups of data mentioned in the procedure?
	Reflection	Asking students to extend their thinking or providing topics that they previously learned.	What methods have we done in class to help visualize how the data is distributed?
Cognitive Statistical	Central tendency	The way in which the values of a random variable cluster around a certain value.	They are both a measure of central tendency.
	Variation	The extent to which data differ from each other in data sets.	However it does not take into account its variability.
	Distribution	Various characteristics of data, such as shape, center, and spread.	You have no really taken the distribution of data into consideration.
	Precision	Addressing mathematical misconception or miscalculation.	You should not remove outliers.
Others	Redundancy	Repetition of the same contents.	Where is the distribution of data . . . Consider the distribution of data.
	Reproduction	Providing a copy from content in I-MAP or course materials without rephrasing it.	Identify the complexity in the problem.

study because these last two feedback characteristics were evident in TAs' feedback.

Feedback coded as summary acknowledges the main points in students' work. Effective summaries reorganize information into more manageable pieces [37]. As part of summary, feedback may also provide information about the results when the instructor applies the mathematical model that the students developed. When an instructor recognizes problems in the students' mathematical model, he or she might provide feedback regarding the incorrect part of the mathematical model. In our study, this extension of summary feedback is coded problem and identifies something that needs to be corrected.

Another component of feedback includes recommendations, which are coded as solutions. Solutions are statements that directly suggest ways to improve the work. Instructors are encouraged to provide specific feedback that offer suggestions for improvement [38]. This type of feedback is slightly different from "reflections" which are statements designed to help students extend their ideas by referring to topics that students have previously experienced. Driscoll [36] described reflections as questions that invite further thoughts on or extend students' rea-

soning about the mathematical concepts in a problem. The following examples may clarify the distinction between solution and reflection.

Solution: Look at a histogram of the data and observe its distribution.

Reflection: What about other statistical measures that could inform the user?

A solution or reflection is sometimes followed by explanations. Without explanation of what to revise or why it should be revised, students may not understand the feedback. When those who provide feedback do not understand what students intend to convey, they request further information or more details from the students. These requests were coded as clarification.

3.3 Cognitive feedback addressing statistical concepts

Statistical cognitive feedback includes central tendency, variation, distribution, and mathematical precision—concepts used in solving the *JITM MEA*. One purpose of the *JITM MEA* is for students to experience a statistically complex pro-

blem and to develop a solution using statistics appropriately. To achieve this goal, TAs were advised to provide feedback on students' use of the statistical concepts in their models via written feedback. According to the *I-MAP*, students were to decide how to interpret uni-variate data. Their solutions needed to go beyond concepts of mean and standard deviation as these data characteristics are not sufficient for differentiating the data sets provided in the problem for decision making purposes in the context of the problem (Appendix B). However, students struggle with when and how to use central tendency, variation, and distribution [39] and mathematical precision.

Central tendency refers to the way in which the values of a random variable cluster around a certain value [40]. Mode, median, and mean are measures of the central value. Variation is also an important data characteristic that students need to consider when they examine data sets [39]. Some variation measures are range, variance, standard deviation, and maximum/minimum. In the *JITM* MEA, students need to move beyond measures of central tendency and variation in the development of their models because the test case data sets are designed such that differences in the mean and standard deviation are not practically significant. Thus, students should consider the distribution of the data provided. The *I-MAP* suggests several aspects of distribution that students could employ in their models: frequency of maximum and minimum values or values within intervals; difference between the mean and the median; and quantification of the shape of the distribution.

Lastly, high quality solutions should be devoid of mathematical and statistical misconceptions. For instance, students often remove outliers from the data even though it is not appropriate in the context of the *JITM* MEA because there is no information to justify the removal of data points. Since this misconception is frequently shown in students' models [25], TAs were guided to provide feedback regarding this concept.

3.4 Other feedback characteristics

The quality of TA feedback is also a factor when analyzing the effect of feedback. Some feedback might include the same statements over and over. The impact of repeated feedback on improving students' understanding is unclear. Previous publications indicated that most repeated feedback statements come from content in the *I-MAP* or other course materials [41, 42]. Even though these statements are copied exactly from related documents, it is uncertain if the feedback influences the students' work. Therefore, analyzing the redundancy and reproduction of TA feedback could

clarify the effect of such feedback characteristics on student work.

4. Methods

4.1 Settings

This study was conducted in the second of a required first-year engineering course sequence focused on problem solving, design, and computer tools at Purdue University. As part of the Fall 2011 and Spring 2012 course sequence, students solved three MEAs. The *JITM* MEA (See Appendix A) was the first MEA solved in the second course, and the second MEA for the year. The *JITM* MEA required teams of three or four students to apply their mathematical and statistical knowledge to develop a mathematical model (Draft 1); then revise their model (solution) twice (Draft 2 and Team Final Response). Revisions were prompted by peer feedback on Draft 1, provision of an expanded data set following Draft 1, and TA feedback on Draft 2.

An appropriate solution consisted of a mathematical procedure or model for solving the problem, rationales for critical decisions, a preface indicating the direct user of the model and what the direct user needs, and the results of applying the model to the test cases (data) provided in the context of the problem. A high quality mathematical model addresses the complexity inherent to or embedded in the problem context. In the context of the *JITM* MEA, a mathematical model is said to be of high quality when there is recognition that measures of central tendency and variation are not sufficient to make decisions with the data provided; and common misconceptions about how data can be interpreted and manipulated are no longer present.

After being instructed on how to give an effective peer review, practicing giving peer feedback on a sample MEA solution, and comparing their feedback to an expert's feedback, students individually provided feedback on their classmates' Draft 1 MEA solutions. Each team received one to four peer reviews to use when revising their Draft 1 before submitting Draft 2. TAs provided teams with written feedback and a current level indicator on their Draft 2. Each team received feedback from only one TA who was either a trained graduate or undergraduate student. After teams revised Draft 2, they submitted their Team Final Response, which was evaluated again by the TA and assigned a final level. Student team work is assigned a level from four (highest) to zero (lowest), which in combination with participation in various parts of the MEA implementation sequence is mapped to an overall numeric grade.

TAs played a significant role in the first-year engineering course sequence because they analyzed student work on MEAs, provided feedback, and gave final grades. TAs participated in five-hours of training per MEA implemented in the first-year engineering courses. The training session provided information about realistic mathematical modeling problems, the process of modeling, and the goals of the courses and MEAs [40]. The process of assessing MEAs and the role of feedback were also discussed. Before, during, and after the training, TAs assessed and provided feedback on prototypical student solutions using the generic, four-dimension (i.e., mathematical model, share-ability, re-usability, and modifiability) *MEA Rubric* and *JITM MEA-specific I-MAP* [43]. A portion of the *JITM MEA I-MAP* describing how TAs were expected to provide feedback and grade student teams' mathematical models is included in Appendix B. The description of the grade levels for the mathematical model dimension indicated the quality of students' solutions. Once TAs completed each assessment, they compared their assessment to that of an expert. This was the primary mechanism for instructing TAs on how to provide constructive feedback. TAs were encouraged to use constructive written feedback to provide information about whether students' solutions were heading in an appropriate direction [24]. The *MEA Rubric*, the *I-MAP*, and the TA training have been designed and evaluated to ensure as consistent an assessment and evaluation of students work as possible across TAs [43, 44].

4.2 Data collection

A web-based MEA management system allowed students to submit their team drafts and managed both the peer and TA feedback processes [45]. Data for this study included students' ratings and explanations on a *Confidence Reflection* (following Draft 2 and Team Final Response), the TA-issued grades on Draft 2 and the Team Final Response, and TA feedback on Draft 2. Draft 2 was the focus of this analysis because students revised their work based upon TA feedback on this draft.

The *Confidence Reflection* consisted of students' reflections on the four *MEA Rubric* assessment dimensions. We focused this analysis on the mathematical model dimension, which determined the extent to which students' solutions addressed the mathematical complexities inherent to the problem.

After submitting Draft 2 and Team Final Response, students completed the *Confidence Reflection* which asked students to individually 1) use a 4-point scale (strongly agree, agree, disagree, or strongly disagree) to rate their confidence that their team's MEA solution addressed the *Mathe-*

matical Model MEA Rubric dimension and 2) supply reasons for their rating. The highest confidence rating, "strongly agree," was assigned four points; the lowest rating, "strongly disagree," was assigned one point.

After student teams submitted their Draft 2 solutions, TAs provided written feedback on the Mathematical Model dimension in three web-form sections: summary, application, and recommendations. This system requires TAs to write their comments in a structured way to help students comprehend the main points of the TA feedback.

4.3 Data analysis

The focus of this study was on identifying characteristics of TA feedback that possibly improved students' mathematical models, and consequently influenced their confidence and MEA grade. To analyze characteristics of TA feedback that influenced student work, we asked questions such as (a) "Did the students mention that they revised their work based on TA feedback?" and (b) "What characteristics of TA feedback seemed to influence student work?"

To answer the first question, we examined students' explanations for their Team Final Response *Confidence Reflection* to find any that explicitly stated their confidence was influenced by TA feedback. Because the total number of students was large ($n = 1,655$), the first author used the "Find" tool to locate explanations that included the terms "feedback," "TA feedback," "expert," "reviewer," "revised," and "modified." Additional terms, such as "revision" or "correction," yielded students' explanations that were not pertinent to this study. At the suggestion of another author, the search term list was expanded to include terms such as "comments," "response," "grader," and "changed." After cycles of improvement, a final list of terms emerged as the most effective for targeting responses referring to TA feedback. The final list of search terms consisted of 13 terms: "feedback", "TA feedback", "comments", "response", "grader", "instructor", "teacher", "expert", "reviewer", "fixed", "changed", "revised", and "modified."

This search yielded a total of 132 *Confidence Reflection* explanations. For example, one student wrote, "After feedback and team discussion, we changed our weights and eliminated some categories to make the model more accurately reflect what the direct user asked for. . ." For a complete example of this student's response, see Appendix C. These 132 reflection explanations came from students on 50 teams. Since the second research question focuses on the identification of the characteristics of influential TA feedback, we ana-

lyzed the TA feedback to these 50 teams using the coding scheme in Table 1.

4.4 Coding process

The TA feedback given to the 50 teams was divided into 680 segments with each segment containing a single idea. Each of the 680 segments was coded, first by affective domain (positive, negative, or neutral), then by general cognitive feedback (summary, problem, solution, explanation, clarification, and reflection), then by statistical cognitive feedback (central tendency, variation, distribution, or mathematical precision), and finally for redundancy or reproduction.

To establish reliability, the coding scheme was applied by two researchers to four complete sets of TA feedback that consisted of a total of 26 segments. Their codes were compared to check inter-rater reliability by using the fixed-marginal *kappa* value [46]. The inter-rater reliability between the coders ranged from Kappa values of 0.76 to 0.85 (Affective: 0.76; Cognitive General: 0.80; Cognitive Statistical: 0.85; and Others: 0.78). According to Brennan and Prediger [46], a *kappa* of 0.70 or above shows adequate inter-rater agreement. The coders discussed each discrepancy in the coded segments until they reached an agreement on all 26 segments. The remaining segments of the TA feedback were then analyzed by the first author.

4.5 Indicating cases of student groups

With the 132 student explanations of their *Confidence Reflection* implying that their confidence was influenced by TA feedback, we analyzed the extent to which these students changed both their confidence and grade (Mathematical Model Dimension level) from Draft 2 to Team Final Response. Table 2 shows the change (if any) in students' confidence and/or grade. For instance, 26 students increased their grade by two or three levels and increased their confidence from Draft 2 to Team Final Response. The confidence rating is categorized as having either increased, stayed the same, or decreased.

Over 92% of students whose confidence increased ($n = 39$), only increased by 1 point. There was no student whose confidence increased over 3 points, though three students' confidence increased by 2

points and six students' confidence decreased by 1 point.

From combinations of math model grade and confidence rating changes, three student confidence-performance categories were identified. Students whose confidence increased and grade increased at least two levels ($n = 26$) were categorized as *substantial increase*. Students whose confidence remained the same and whose grade increased only 1 level ($n = 17$) were categorized as *some increase*. Students categorized as *no increase* included students whose confidence and grade both stayed the same or decreased ($n = 7$). These students did not change their confidence and performance, or they had overconfidence in their performance. All of the students whose grade remained the same ($n = 7$) were at level 2 for Draft 2 and Team Final Response regardless of their change in confidence. The group that increased in grade but had the same confidence was not counted as a case because approximately 82% of the group already had the highest confidence level (4 points) before they received TA feedback. By categorizing three distinct cases, it is possible to compare the characteristics of TA feedback that are likely to be effective for students to improve their work. For example, if students in the *substantial increase* category received more explanatory feedback than those in the *some increase*, while students in the *some increase* category received more explanatory feedback than those in the *no increase*, it is possible that explanatory feedback might be effective for improving student performance and confidence.

4.6 Indicating two teams' use of feedback and changes in confidence

We did a case analysis of two teams to show how different characteristics of feedback influenced students' model development and impacted students' confidence. Two teams were chosen using two criteria: the change in their grade for the mathematical model dimension and the change in their confidence ratings. Team A, from the *substantial increase* category, increased their mathematical model dimension grade from level 1 on Draft 2 to level 3 on Team Final Response. Team B, from the *no increase* category, maintained their grade at level 2. Three of the four students from Team A increased

Table 2. Changes in student confidence and mathematical model dimension grade from Draft 2 to Team Final Response ($n = 132$)

		Confidence Rating		
		Increased	Same	Decreased
Math Model Dimension	Increased (2 or 3 levels)	26	65	4
Grade	Increased (1 level)	11	17	0
	Same or Decreased	2	5	2

Table 3. Number of feedback segments for each feedback type and characteristics, and team category

Type	Characteristics	Substantial Increase ($n = 26$)	Some Increase ($n = 17$)	No Increase ($n = 7$)
Affective	Total # of segments	304	257	119
	Positive	24 (7.9%)	10 (3.9%)	3 (2.5%)
	Negative	0 (0%)	7 (2.7%)	9 (7.6%)
	Neutral	280 (92.1%)	240 (93.4%)	107 (89.9%)
Cognitive-General	Total # of segments	364	286	130
	Summary	109 (29.9%)	92 (32.2%)	29 (22.3%)
	Problem	59 (16.2%)	52 (18.2%)	40 (30.8%)
	Solution	60 (16.5%)	44 (15.4%)	26 (20.0%)
	Explanation	72 (19.8%)	61 (21.3%)	21 (16.15%)
	Clarification	5 (1.4%)	14 (4.9%)	8 (6.15%)
	Reflection	59 (16.2%)	23 (8.0%)	6 (4.6%)
Cognitive-Statistical	Total # of segments	210	131	63
	Central tendency	41 (19.5%)	28 (21.4%)	11 (17.5%)
	Variation	57 (27.1%)	41 (31.3%)	13 (20.6%)
	Distribution	101 (48.1%)	46 (35.1%)	33 (52.4%)
	Precision	11 (5.2%)	16 (12.2%)	6 (9.5%)
Other	Total # of segments	304	257	119
	Redundancy	5 (1.7%)	12 (4.7%)	16 (13.5%)
	Reproduction	0 (0%)	0 (0%)	7 (5.9%)

their confidence ratings; none of the Team B students increased their confidence ratings.

5. Findings

5.1 Types of feedback students received

The first research question focused on identifying the types of feedback that possibly influenced students' confidence and performance. In this section, we describe examples of feedback given to teams in the three categories (i.e., *substantial increase*, *some increase*, and *no increase*) as we focus on four feedback types: affective, general, statistical, and other. Table 3 summarizes all of the results including the number of feedback segments and percent for each feedback type for the three team categories.

5.1.1 Affective domain

In the team categories shown in Table 3, neutral segments accounted for the greatest percentage of feedback in the affective domain. While most feedback segments were neutral, teams in the *substantial increase* category received more positive feedback (7.9%) than those in the *some increase* category (3.9%). Similarly, teams in the *some increase* category received more positive feedback than those in the *no increase* category (2.5%). Some of the positive feedback that teams received included "It was good to include frequency in your analysis" (Team 1), "I like the analysis of 'percent on time'" (Team 2), and "Using a histogram is a good start, and gives a good indicator of the distribution of the data" (Team 3). With respect to negative feedback, teams in the *substantial increase* category received no negative feedback, while those in the *no increase* category

received more negative feedback (7.6%) than those in the *some increase* (2.7%) category. Some of the negative feedback included "Don't just assume your reader will agree" (Team 4), and "That whole idea is very confusing" (Team 5). As shown in Table 3, students who improved their confidence and performance (*substantial increase*) tended to receive more positive and less negative feedback overall.

Since one positive or negative segment in a TA's feedback might influence the motivation of students, we also considered the number of teams in each confidence and performance category who received at least one positive or negative feedback segment (Table 4). At least one third of all teams in each category received TA feedback that included positive segment(s). The positive feedback was mostly given with additional suggestions to improve student work. About 44% of the teams in the *no increase* category received negative feedback while teams in the *substantial increase* category did not receive any negatively phrased feedback.

5.1.2 General cognitive feedback

With regards to the receipt of general cognitive feedback, patterns were detected for the feedback types of problem, reflection, and clarification. For

Table 4. Number of teams receiving at least one affective feedback segment

Category	Substantial Increase ($n = 26$)	Some Increase ($n = 17$)	No Increase ($n = 7$)
Positive	15 (57.7%)	9 (52.9%)	3 (33.3%)
Negative	0 (0%)	2 (11.8%)	4 (44.4%)
Neutral	26 (100%)	17 (100%)	7 (100%)

example, teams in the *no increase* category received the greatest percent of problem feedback (30.8%), while those in the *some increase* and *substantially increase* categories received lower percentages of this type of feedback (18.2%, 16.2%, respectively). That is, students whose confidence and grade increased received fewer feedback comments identifying what students needed to fix than the other students. It is unlikely that *substantial increase* students' solutions were less problematic than that of *some increase* or *no increase* because all of their initial grades (the grades on Draft 2) were only at level 1 or 2. Some of the examples of feedback indicating problems in teams' solutions were "This does not give a clear and viable way of ranking" (Team 6), "I did have a slight issue in applying your procedure" (Team 7), "You are not getting the correct results" (Team 8), and "Making the histogram is somewhat unnecessary for this procedure" (Team 9).

Instead of receiving feedback indicating problems in their solutions, *substantial increase* teams received more reflective feedback (16.2%) than *some increase* teams (8.0%), and over three times more than *no increase* teams (4.6%). Some examples of reflective feedback comments included "How will you rank the companies if two or more of them have the same value at the end of the calculation?" (Team 10), "Also what would happen if 2 or more companies have the same total number of points?" (Team 11), "It was explicitly stated in class that you are to use all the data given to you . . . What methods have we learned to visualize the distribution of a dataset? Also look at what the numbers actually mean and how much they might actually matter" (Team 12), and "What if there is a company that has shorter late times compared to another company which has an even distribution of late times?" (Team 13). The results show that students who increased their performance and confidence received feedback including topics related to students' prior knowledge or including feedback that helped expand their current thinking.

Another notable pattern is that teams in the *substantial increase* and *some increase* categories received the lowest percent of clarification feedback. Specifically, *substantial increase* teams received fewer clarification feedback segments (1.4%) than *some increase* teams (4.9%) and *no increase* teams (6.2%). Some of the clarification feedback included "What are these groups? . . . What are these groups of data mentioned in the procedure? How do I find them?" (Team 14), "Is your plot a histogram?" (Team 15), and "Did you plot the histograms?" (Team 16). When TAs provided feedback including clarification, it might have been confusing to students because it is not easy to interpret a TA's actual

intention of asking such questions in the online environment.

5.1.3 Statistical cognitive feedback

A noticeable pattern is that teams in all three categories received the greatest percent of feedback on distribution (48.1% for *substantially increase*, 35.1% for *some increase*, 52.4% for *no increase*), followed by variation (27.1%, 31.3%, 20.6%, respectively), central tendency (19.5%, 21.4%, 17.5%, respectively), and precision (5.2%, 12.2%, 9.5%, respectively), in that order. Teams received feedback on distribution including, "So it is important that you take a look at the distribution of the data and decide on how to incorporate that into your procedure" (Team 6), "Also, you have no [not] really taken the distribution of data into consideration" (Team 18), "In case of a tie between frequencies of perfect deliveries, companies with a higher frequency of delay times between one and four hours, were ranked higher than those with lower frequencies in the 0 to 4 hours bin" (Team 19), and "Consider the data you are looking at, it is very skewed and not much like the normal distributions you are used to dealing with" (Team 20).

Teams also received feedback on variation, such as "Your procedure takes into account the variability of the data" (Team 6), "What are other measures of variability?" (Team 1), "Your model currently only accounts for the variability of the data" (Team 21), and "Almost all the companies fall between the 1.5ish range" (Team 22). In terms of central tendency, teams received feedback including, "The use of average to determine the ranks is insufficient on its own" (Team 23), "Mean was calculate for the data sets. Lower the mean better the company's ranking. . . How far apart are the means?" (Team 24), and "Some of your mathematical means of calculating like MODE is not a good way to differentiate one company from another, if you realize 4 companies have mode of 0, this would actually reduce the differentiating factor in calculating total points" (Team 22).

Lastly, teams received feedback on precision, such as "Don't take out outliers" (Team 12) and "Reconsider what eliminating the 0s really does" (Team 25). All of this feedback seemed to help teams revise their mathematical models because this feedback directly indicated statistical problems in teams' current models and suggested possible ways to improve the models by focusing on their use of statistical measures.

5.1.4 Redundancy and reproduction

As shown in Table 3, teams in the *no increase* category (13.5%) received more redundant segments than those in the *some increase* category

(4.7%), while teams in the *some increase* category received such segments more often than those in the *substantial increase* category (1.7%). For example, a team in *no increase category* received feedback including “Unable to generate results. Most of your procedure is confusing and hard to understand. Therefore, I was unable to generate any results . . . The procedure needs to be rewritten to address the task at hand . . . See feedback from above. Write concise steps that are easy to follow . . . Need to be specific (Team 14).” This feedback included redundant comments; if it was more simply written, the meaning would be “unable to generate results, so need to write concise and specific steps.”

In terms of reproduction, only students in the *no increase* category (5.9%) received feedback reproduced directly from the MEA *I-MAP*. Such feedback included “Identify the complexity in the problem” (Team 26) and “Develop a model that is simple and elegant but addresses the complexity of the problem” (Team 26). This feedback is general and does not indicate to the students what and how to revise their models.

5.2 Two teams’ use of feedback and changes in confidence and model

While the previous section describes different types of feedback that students received, this section uses the work of two teams (i.e., Team A and B) to investigate how these different types of feedback impacted students’ confidence and model development. Team A was in the *substantial increase* category whereas Team B was in the *no increase* category. For Draft 2, Team A’s model only employed the mean of the data without appropriately addressing the variation in or distribution of the data (Level 1). Team A improved their model for the Team Final Response by addressing both the variation in and distribution of the data, as well as including a ranking procedure (Level 3). On the other hand, Team B maintained their mathematical model level from Draft 2 to Team Final Response. They considered variation, but did not address distribution of the data. Their mathematical detail was lacking and mathematical errors were still present in their Team Final Response (Level 2).

5.2.1 Affective domain

Team A received feedback including several positive but no negative feedback segments. For example, the TA feedback included such statements as, “The percentage of on-time deliveries is a good way to rank shipping companies” and “The first tie-breaking method is a good way.” The team kept the percentage of on-time deliveries and tie-breaking methods in their Team Final Response. Their

explanations for their Team Final Response Confidence Rating showed these positive comments influenced their confidence. Student 1 wrote, “I believe we cover the problem well and have adequate tie breakers if such an occasion arises.” Student 2 also mentioned, “We clearly state the mathematical meaning of the value of percentage of the deliveries within four hours or within one standard deviation in this case.”

Team B, on the other hand, received several negative but no positive feedback segments, even though Team B’s solution in Draft 2 better addressed the complexity of the problem than Team A’s solution. TA feedback included such statements as, “Most of your procedure is confusing and hard to understand,” and “I was unable to generate any results.” There was no evidence shown in the students’ confidence reflections that these negative comments affected their confidence. One of the students wrote, “Our team’s solution addresses the mathematical complexity of the problem because it uses three separate computations to compare the data sets.” The team’s Final Response shows, however, that they did not improve their procedure because their three computations included average late time and the standard deviation, which was the same as in their Draft 2.

5.2.2 Other domains

Team A received feedback that included several segments of explanation, reflection, and distribution, but no clarification or redundancy segments. TA feedback included segments such as, “However, looking at the results (very less difference in percentages of on-time deliveries for each company), is it a good criteria to rank shipping companies? Think about standard deviation, mean and mode and about what these statistical features signify and how they can be used in your solution.” This comment is actually followed by the positive comment mentioned earlier, “The percentage of on-time deliveries is a good way to rank shipping companies.” By asking explanatory and reflective question, the TA feedback changed the positive feedback to mitigation. Student 4 mentioned in his or her confidence reflection, “Based on the grader reflection from the Draft 2, we made a lot of improvement. We change the hard coded values (4 hours late) to use the frequency of those late deliveries within one standard deviation of delivery times of these companies as our first tiebreaker.”

Team B, however, received feedback segments including problem, clarification, and redundancy, but fewer distribution segments, even though the team did not adequately use distribution in their mathematical model. For example, the TA feedback included such statements as, “Why are histograms

generated. . . they are not used in the procedure to help address the problem,” and “The procedure needs to be rewritten to address the task at hand.” These feedback comments, including problem segments, were repeated, “Your team is generating histograms and are not being used in the procedure,” and “See feedback from above. Write concise steps that are easy to follow.” One of the students in Team B wrote, “We completely redid our model after the grader commented on it. Now the model incorporates the sample means, sample standard deviations, and the sample percentage of late arrivals.” As the students mentioned, the team did not address distribution in their mathematical model in the Team Final Response. The team took the TA’s comment about not using distribution as an indicator to drop it out altogether. The TA might have been more effective if the feedback indicated that looking at distribution is good and should be incorporated into the model.

6. Discussion

The participants in this study were students whose initial confidence levels were fairly high. Kulhavy [18] contended that students might not pay much attention to feedback about a task when their initial confidence is low because they might not find the feedback helpful. When they initially struggle with the nature of the task, feedback consisting of a mixture of students’ existing knowledge and new information may not help them. On the other hand, students pay special attention to feedback when their confidence level is high and their answer to the task is wrong, because they are willing to improve their solution. The participants in this study seem to pay attention to feedback because their initial confidence levels were high but their grades on the task were low. Nevertheless, Carnes et al. [25] reported that students seemed to make only minor changes after getting TA feedback; as such, they recommended further examination of the nature of TA feedback. Since students’ limited use of feedback has already been studied at length, we focused on the TA feedback that was given to students who made gains in performance and confidence, along with the feedback that was given to students who did not make gains, to inform other educators of the nature of feedback.

In the affective domain, we found that students who increased their performance and confidence tended to receive more positive feedback, such as praise. Whether or not providing praise positively influences learning is still controversial [47]. Our finding is contrary to others’ research findings that caution against the practice of providing praise, which suggested that praise to learners could dis-

tract their attention from the task [35, 48]. While it is possible that learners may not improve their work after receiving praise alone, in this study, praise was mostly given with additional suggestions to improve student work. The question, then, is how to provide praise in a manner that increases learners’ intrinsic motivation and confidence without distracting their attention from the task. Praise, when perceived as sincere, can encourage autonomy and improve competence [47]. Furthermore, praise can be regarded as a motivator that increases students’ willingness to revise their work [48, 49]. At the same time, negative feedback that focuses on learners, rather than the task, is generally conceived of as ineffective feedback in other studies [7] and this was supported by this study.

For the general cognitive feedback component, students consistently increased their confidence and performance when they received less feedback indicating problems or requesting clarifications and more feedback encouraging reflection. Feedback identifying what students missed or needed to fix is not effective without knowing how to improve it [33]. Similarly, clarification statements did not suggest ways students should improve their solution. As Shute [7] suggested, feedback should be specific and clear by providing goals of the task and anticipated performance. Reflective feedback meets such conditions because it provides students with topics that they previously learned or includes questions that extend their thinking.

In this study, the number of other characteristics of general cognitive feedback, summary, solution, and explanation, did not result in differences in students’ performance or confidence. This result is different than that presented by Nelson and Schunn [33]; they concluded that solutions and summarization support students’ understanding of the problem. The exact reasons for the conflicting results between Nelson and Schunn’s study and this study are uncertain. It could be due to the different natures of the tasks (historical writing vs. mathematical modeling), sources of the feedback (peer review vs. TA feedback), foci of the studies (performance only vs. performance and confidence), or prior knowledge of the participants.

For the statistical cognitive feedback component, the greatest percent of the feedback related to distribution concepts. Carnes et al. [25] found that students employ distribution less often than other statistical concepts to solve the *JITM MEA*, though an analysis of the data using distribution is necessary to distinguish between the shipping companies. So, the *I-MAP* leads TAs to guide students towards looking at distribution. Shaughnessy [39] also claims that distributional reasoning is one step further along since it includes both central tendency

and variability concepts. Shaughnessy's claim is connected to why distribution appeared so much in the TA feedback in this study. Further efforts to help students understand distribution concepts are necessary and effective feedback addressing distribution might be one way to do it.

Finally, students who increased their confidence and performance received less redundant feedback and less feedback that is copied from other resources. Thus, it would be more helpful to keep feedback simple, rather than provide the same content over and over. Also, providing feedback that is directly related to learners' solutions is more effective than providing generic feedback from a rubric.

7. Conclusions

The aim of this study was to describe characteristics of written feedback that possibly influenced students' confidence and performance in addressing the complexity embedded in an authentic mathematical modeling problem and to show how different types of written feedback impacted students' confidence and model development. We described characteristics of affective and cognitive written feedback given to first-year engineering students. The findings show that students who substantially improved their model and increased their confidence received more positive and reflective feedback than those who received redundant and generic feedback or feedback that indicates problems and requests clarifications. A limitation is that, within the current data collection and analysis, it cannot be known the extent to which other factors, such as students' prior knowledge, may be contributing to or detracting from student confidence. Future studies might consider how students' previous knowledge influences the nature of TA feedback and students' model development. The specific characteristics of feedback described in this study can be useful for instructors who wish to implement mathematical modeling problems with an iterative process of feedback and revision. Ultimately, it is evident that students improve their performance and confidence in developing mathematical models if they are in an environment where they receive effective feedback to help them revise their work.

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

References

1. R. L. Bangert-Drowns, C. C. Kulik, J. A. Kulik and M. T. Morgan, The instructional effect of feedback in test-like

- events. *Review of Educational Research*, **61**, 1991, pp. 213–238.
2. R. Moreno, Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia, *Instructional Science*, **32**, 2004, pp. 99–113.
3. D. R. Pridemore and J. D. Klein, Control of practice and level of feedback in computer-based instruction, *Contemporary Educational Psychology*, **20**, 1995, pp. 444–450.
4. A. P. Volpentesta, S. Ammirato and F. Sofo, Collaborative design learning and thinking style awareness, *International Journal of Engineering Education*, **28**(4), 2012, pp. 948–958.
5. S. Narciss and K. Huth, How to design informative tutoring feedback for multimedia learning, In H. M. Niegemann, D. Leutner & R. Brunken (Eds.), *Instructional design for multimedia learning*, 2004, pp. 181–195. Munster, NY: Waxmann.
6. C. J. Shieh, S. Zhifang and S. P. Yeh, Key success factors in cultivating students' learning motivation, *International Journal of Engineering Education*, **30**(2), 2014, pp. 326–332.
7. V. J. Shute, Focus on formative feedback, *Review of Educational Research*, **78**(1), 2008, pp. 153–189.
8. M. Briggs, Models and modeling: a theory of learning. In G. M. Bodner & M. K. Orgill (Eds.), *Theoretical frameworks for research in chemistry/science education*, 2007, (pp. 72–85). NJ: Prentice Hall.
9. R. Lesh, M. Hoover and A. E. Kelly, Equity, assessment, and thinking mathematically: Principles for the design of model eliciting activities, In I. Wirszup & R. Streit (Eds.), *Developments in school mathematics education around the world* (Vol. 3), 1993, (pp. 104–130), Reston, VA: National Council of Teachers of Mathematics.
10. W. Z. Bernstein, D. Ramanujan, F. Zhao, K. Ramani and M. F. Fox, Teaching design for environment through critique within a project-based product design course, *International Journal of Engineering Education*, **28**(4), 2012, pp. 799–810.
11. G. W. Ellis, A. N. Rudnitsky, M. A. Moriarty and B. Mikic, Applying knowledge building in an engineering class: a pilot study, *International Journal of Engineering Education*, **27**(5), 2011, pp. 945–957.
12. R. Lesh, M. Hoover, B. Hole, A. Kelly and T. Post, Principles for developing thought-revealing activities for students and teachers, In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of Research Design in Mathematics and Science Education*, 2000, (pp. 591–645), Mahwah, NJ: Lawrence Erlbaum.
13. National Council of Teachers of Mathematics, *Principles and standards for school mathematics*, Reston, VA: National Council of Teachers of Mathematics, 2000.
14. National Governor's Association Center for Best Practices & Council of Chief State School Officers. *Common core state standards for mathematics*. Washington, DC: National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010.
15. Accreditation Board of Engineering and Technology Accreditation Department, http://www.abet.org/uploadedFiles/Accreditation/Accreditation_Step_by_Step/Accreditation_Documents/Current/2013_-_2014/eac-criteria-2013-2014.pdf, Accessed 4 July, 2014.
16. R. Bénabou and J. Tirole, Self-confidence and personal motivation, *Quarterly Journal of Economics*, **117**(3), 2002, pp. 871–915.
17. R. Clément, Z. Dornyei and K. A. Noels, Motivation, self-confidence, and group cohesion in the foreign language, *Language Learning*, **44** (3), 1994, pp. 417–448.
18. R. W. Kulhavy, Feedback in written instruction, *Review of Educational Research*, **47**(2), 1977, pp. 211–232.
19. H. A. Diefes-Dux and P. K. Imbrie, Model-Eliciting Activities in a first-year engineering course, In J. S. Zawojewski, H. A. Diefes-Dux & K. Bowman (Eds.), *Models and modeling in engineering education: designing experiences for all students*, 2008, (pp. 55–92), Rotterdam, the Netherlands: Sense Publishers.
20. L. English and R. Lesh, R. Ends-in-view problems, In R. Lesh and H. Doerr (Eds.), *Beyond constructivism: A models and modeling perspectives on mathematics problem solving*,

- learning, and teaching, 2003, (pp. 297–316), Mahwah, NJ: Lawrence Erlbaum Associates.
21. J. S. Zawojewski and R. Lesh, A models and modeling perspective on problem solving, In R. Lesh, and H. Doerr (Eds.), *Beyond constructivism: A models and modeling perspectives on mathematics problem solving, learning, and teaching*, 2003, (pp. 317–336), Mahwah, NJ: Lawrence Erlbaum Associates.
 22. K. J. Rodgers, A. K. Horvath, H. Jung, A. S. Fry, H. A. Diefes-Dux and M. E. Cardella, Students' perceptions of and responses to teaching assistant and peer feedback, *Interdisciplinary Journal of Problem-based Learning*, **9**(2), 2014, <http://dx.doi.org/10.7771/1541-5015.1479>.
 23. M. T. Carnes, H. A. Diefes-Dux and M. E. Cardella, Evaluating student responses to open-ended problems involving iterative solution development in Model Eliciting Activities, 2011, *Proceedings of the 118th ASEE Annual Conference & Exposition, Vancouver, B.C.*
 24. H. A. Diefes-Dux, J. S. Zawojewski, M. A. Hjalmarson and M. E. Cardella, A framework for analyzing feedback in a formative assessment system for mathematical modeling problems, *Journal of Engineering Education*, **101**(2), 2012, pp. 375–406.
 25. M. T. Carnes, M. E. Cardella and H. A. Diefes-Dux, Progression of student solutions over the course of a model-eliciting activity (MEA). *Proceedings of the 40th ASEE/IEEE Frontiers in Education Conference*, 2010, Washington, DC.
 26. J. Hattie and H. Timperley, The power of feedback, *Review of Educational Research*, **77**(1), 2007, pp. 81–112.
 27. R. Slavin, *Education for all*, Exton, PA: Swets & Zeitlinger Publishers, 1996.
 28. A. S. Palincsar and A. L. Brown, Reciprocal teaching of comprehension fostering and comprehension monitoring, *Cognition and Instruction*, **1**(2), 1984, pp. 117–175.
 29. L. S. Vygotsky, Interaction between learning and development (M. Lopez-Morillas, Trans.), In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in society: The development of higher psychological processes*, 1978, (pp. 79–91), Cambridge, MA: Harvard University Press.
 30. W. H. Melching, Programmed instruction under a feedback schedule, *National Society of Programmed Instruction Journal*, **5**(2), 1966, pp. 14–15.
 31. R. W. Kulhavy and W. Wager, Feedback in programmed instruction: Historical context and implications for practice, In J. V. Dempsey & G. C. Sales (Eds.), *Interactive instruction and feedback*, 1993, (pp. 3–20), Englewood Cliffs, NJ: Educational Technology.
 32. D. R. Ferris, The influence of teacher commentary on student revision, *TESOL Quarterly*, **31**(2), 1977, pp. 315–339.
 33. M. M. Nelson and C. D. Schunn, The nature of feedback: How different types of peer feedback affect writing performance, *Instructional Science*, **27**(4), 2009, pp. 375–401.
 34. S. C. Tseng and C. C. Tsai, On-line peer assessment and the role of the peer feedback: A study of high school computer course, *Computers & Education*, **49**(4), 2006, pp. 1161–1174.
 35. A. N. Kluger and A. DeNisi, The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory, *Psychological Bulletin*, **119**(2), 1996, pp. 254–284.
 36. M. Driscoll, *Fostering Algebraic Thinking: A Guide for Teachers, Grades 6–10*. Heinemann, 361 Hanover Street, Portsmouth, NH 03801-3912, 1999.
 37. J. D. Bransford, A. L. Brown and R. R. Cocking, *How people learn: Brain, mind, experience, and school* (Rev. ed.), Washington, DC: National Academies Press, 2000.
 38. L. Jennifer, M. Trevisan, J. McCormack, S. Beyerlein, D. Davis, P. Leiffer, P. Thompson, H. Davis, S. Howe, P. Brackin, R. Gerlick and M. J. Khan, Alumni perspective on professional skills gained through integrated assessment and learning, *International Journal of Engineering Education*, **30**(1), 2014, pp. 48–59.
 39. M. Shaughnessy, What do we know about students' thinking and reasoning about validity in data? *Student Learning Research Brief, National Council of Teachers of Mathematics*. http://www.nctm.org/uploadedFiles/Research_News_and_Advocacy/Research/Clips_and_Briefs/Research_brief_11_-_Variability.pdf, Accessed 4 July 2014.
 40. Y. Dodge, D. Cox, D. Commenges, P. Solomon and S. Wilson, *The Oxford dictionary of statistical terms*, Oxford University Press, 2003.
 41. H. A. Diefes-Dux, K. Osburn, B. M. Capobianco, and T. Wood, Chapter 12: 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*, 2008b, (pp. 225–255), Rotterdam, the Netherlands: Sense Publishers.
 42. K. Rodgers, H. A. Diefes-Dux, H. Jung and M. E. Cardella, A comparative analysis of feedback from undergraduate and graduate teaching assistants on open-ended problems, *Paper presented at the American Educational Research Association (AERA)*, 2013, San Francisco, CA.
 43. H. A. Diefes-Dux, J. S. Zawojewski and M. A. Hjalmarson, Using educational research in the design of evaluation tools for open-ended problems, *International Journal of Engineering Education*, **26**(4), 2010, pp. 807–819.
 44. M. A. Verleger and H. A. Diefes-Dux, A teaching assistant training protocol for improving feedback on open-ended engineering problems in large classes, *Proceedings of the 120th ASEE Annual Conference and Exposition*, 2013, Atlanta, GA.
 45. M. A. Verleger and H. A. Diefes-Dux, H., Facilitating teaching and research on open-ended problem solving through the development of a dynamic computer tool, *Proceedings of the 117th ASEE Annual Conference and Exposition*, 2010 Louisville, KY.
 46. R. L. Brennan and D. J. Prediger, Coefficient Kappa: Some uses, misuses, and alternatives, *Educational and Psychological Measurement*, **41**(3), 1981, pp. 687–699.
 47. J. Henderlong and M. Lepper, The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin*, **128**(5), 2002, pp. 774–795.
 48. R. Butler, Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance, *Journal of Educational Psychology*, **79**(4), 1987, pp. 474–482.
 49. P. M. Symonds and D. H. Chase, Practice vs. motivation, *Journal of Education Psychology*, **20**(1), 1929, pp. 19–35.

Appendix A. Just-in-Time Manufacturing (JITM) MEA (Spring 2012)

Model-Eliciting Activity (MEA) 1 - Just-In-Time Manufacturing

Interoffice Memo

To:
Applications Engineering Team
From: Devon Dalton, CEO
RE: Shipping Issues
Priority: [Urgent]

Our company operates a just-in-time manufacturing system. After several years of shipping with Pathways Transit (PT), it has come to my attention that PT has not been meeting our shipping needs. We are having problems with late arrival times. The fact that PT is not consistently arriving at the time they have promised is causing D. Dalton Technologies (DDT) production problems. This means that our Logistics Manager needs a method to identify a new shipping company.

I want to make use of your team’s analytical expertise. DDT is small; therefore, we need your team to serve in an engineering project management function on this project. Your team’s task is to design a procedure to rank potential shipping companies. My assistant has collected historical data on several potential companies for you. Eight shipping companies have been identified as able to transport materials directly from Ceramica to Bowman. As you know, arrival time of materials is a big issue for DDT. Since the piezoelectric materials are designed specifically for each custom order, it is imperative that the delivery of materials occur just-in-time for Bowman to begin the manufacturing process that uses all of the shipped materials. Because we operate with a small workforce and only one shift, minutes to a few hours can make a difference in our ability to complete devices for our custom applications by our contracted delivery date. This makes arrival time of materials of great importance. We have in excess of 250 data points for each shipping company. At this time, the data for only four companies is available. This data is stored in a file called *jit_data_partial.txt*. The four shipping companies are Iron Horse Expeditors (IHE), Delphi Shipping (DS), ShipCorp (SC), and United Express (UE). The data is in hours late for shipping runs from Lincoln, Nebraska to Noblesville, Indiana.

Your team should brainstorm different ways in which to analyze the shipping data. Then, your engineering team will use the sampling of data provided for the four shipping companies to develop a procedure to rank the shipping companies in order of most likely to least likely able to meet our timing needs.

In a memo to my attention, please include your team’s procedure and the rank order of the shipping companies generated by applying your procedure to the sample data. Be sure to include additional quantitative results as appropriate to demonstrate the functionality of your procedure. Please be sure to include your team’s reasoning for the each step, heuristic (i.e. rule), or consideration in your team’s procedure.

Please send your complete memo to me by next week.

DD

Sample Data Set from *jit_data_partial.txt*: Number of hours late for shipping runs from Lincoln, NE to Noblesville, IN. (1000 km apart)

Note: The partial data set included 255 data points for each of the 4 shipping companies. A full data set, provided after Draft 1, included 8 shipping companies each with 255 data points.

IHE ^a	DS	SC	UE
0	1.00	0	1.11
1.31	0	7.39	0.90
0	10.49	1.81	0
0	0.70	9.00	1.11
1.73	0.71	4.22	0.84
1.92	0.42	0.32	3.31

^a IHE = Iron Horse Expeditors; DS = Delphi Shipping; SC = ShipCorp; UE = United Express

Appendix B. Mathematical Model Assessment Guideline from *I-MAP* for *JITM* MEA

Levels	Criteria
0	There is no procedure to rank the shipping companies. Merely computing a series of statistical measures without a coherent procedure to use the results falls into this level.
1	The procedure described does not account for either 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.
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.

-
- 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. HOWEVER, distribution is NOT part of the main procedure – it is used to break ties or provide a check; the use of distribution is almost an afterthought.
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.
- 4 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 accounting for distribution is included in the main procedure.
The procedure provides a viable strategy for how to break tie.
Mathematical detail should be clear from start to finish.
Mathematical errors should be eliminated.
-

Appendix C. Sample Student Team Final Confidence Reflection with Search Terms Italicized

Confidence Reflection on Team Final by a student

I am very confident that my team's mathematical model addresses the complexity of the problem. The design of our model, using a weighted decision matrix, was always a good predictor and producing accurate rankings for the given data.

However, we originally had weightings of the categories that favored consistency over timeliness so was not producing the right result. After *feedback* and team discussion, we *changed* our weights and eliminated some categories to make the model more accurately reflect what the direct user asked for in his original memo.

I am now confident that the model can accurately predict the rankings of companies given data sets for lateness of a company.

Hyunyi Jung is a doctoral student in Mathematics Education and an instructor of elementary and secondary mathematics education courses at Purdue University. She has taught mathematics at the elementary, middle, and high school levels in the U.S. and in Korea. She received her M.S. degree in Mathematics Education from Indiana University Bloomington while participating in the Secondary Transition to Teaching Program. Her experiences include a university supervisor and a CCSSM Curriculum Alignment Tool Reviewer. Her research interests center around mathematical modeling, teacher education, and cross-cultural studies.

Heidi A. Diefes-Dux is a 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, teaching and guiding the design of one of the required first-year engineering courses that engages students in open-ended problem solving and design. Her research focuses on the development, implementation, and assessment of model-eliciting activities with authentic engineering contexts. She is currently a member of the educational team for the Network for Computational Nanotechnology (NCN, nanoHUB.org).

Aladar K. Horvath is an Assistant Professor of mathematics at Ivy Tech Community College. He received his B.S. in Mathematics from Purdue University, his M.S. in Mathematics from Michigan State University, and his Ph.D. in Mathematics Education from Michigan State University. He also completed a year-long post-doctoral research appointment in the School of Engineering Education at Purdue University.

Kelsey Joy Rodgers is a doctoral student in Engineering Education at Purdue University. She received her B.S. in Engineering at Arizona State University. She began conducting research within the First-Year Engineering Program in fall 2011. Her research interests center around how students give and receive feedback on mathematical modeling problems, how students develop mathematical models and simulations, and students' understandings of mathematical models and simulations as engineering tools. She is currently a graduate student on the educational team for the Network for Computational Nanotechnology (NCN, nanoHUB.org).

Monica E. Cardella is an Associate Professor in the School of Engineering Education at Purdue University. She received her B.S. in Mathematics from the University of Puget Sound and her M.S. and Ph.D. in Industrial Engineering from the University of Washington. She is a member of Purdue's Teaching Academy. She has been engaged in teaching first-year engineering and guiding the design of the first required first-year engineering course that engages students in mathematical modeling, design and other forms of problem solving since 2007. Her research focuses on the development of engineering problem solving and design skills from preschool to professional practice, particularly in out-of-school settings. Finally, she is the Director for Purdue's Institute for P-12 Engineering Research and Learning (INSPIRE, inspire-purdue.org).