First-Year Engineering Team Responses to Feedback on Their Mathematical Models - A Video Study

Oguz Hanoglu, Purdue University, West Lafayette

Oguz Hanoglu is currently a graduate student at Purdue University in the School of Engineering Education. He received his B.S. in Electrical Engineering from Middle East Technical University (METU), Turkey and M.S. in Electrical Engineering from Bilkent University, Turkey. He conducts research within the First-Year Engineering Program to help the development, implementation, and assessment of model-eliciting activities with authentic engineering contexts. He is also a member of the Network for Computational Nanotechnology (NCN) education team.

Aladar Horvath, Ivy Tech Community College
Prof. Heidi A. Diefes-Dux, Purdue University, West Lafayette

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 the Director of Teacher Professional Development for the Institute for P-12 Engineering Research and Learning (INSPIRE) and a member of the educational team for the Network for Computational Nanotechnology (NCN).
First-Year Engineering Team Responses to Feedback on Their Mathematical Models - A Video Study

Abstract

High quality formative feedback is an integral component of a fruitful learning experience. Pedagogical approaches are needed to increase the quality of instructor feedback and train students to interpret and appropriately respond to feedback. To develop research-informed approaches, students’ thinking from receipt of feedback to action can be explored through an analysis of documented works, such as written feedback and students’ iterative solutions. However, such approaches do not reveal the whole story of their interactions with feedback. The purpose of this study is to explore student team responses to teaching assistants’ (TAs’) written feedback while revising their mathematical model. The research question that guides this study is: How do student teams respond to feedback and convey their ideas from their team discussions in their documented works? We report case findings from two first-year engineering student teams’ responses to TA feedback on a Model-Eliciting Activity (MEA). The teams were videotaped while working to revise their draft. The findings from this data are supported by documented works (written feedback and students’ iterative solutions) along with student interviews. In trying to understand the complexity of students’ learning experience, this study provides insights into how students respond to TA feedback, specifically how they interpret feedback, budget time, and effectively report the outcomes of team discussion. Moreover, findings imply that TAs need to better identify misconceptions and target feedback appropriately.

I. Introduction

Leadership coach Rick Tate said "feedback is the breakfast of champions". This is especially true for education. Students benefit from timely and effective feedback. However, achieving timely and effective feedback is not easy, especially when open-ended problem-solving activities are considered. These activities may require complex product development that involves deep and multidimensional thinking. In open-ended activities, there is not one right answer and in most cases, this will result in a variety of different answers from students. The challenge here is to provide timely feedback by reviewing all answers in detail, deciding on an effective way to guide each unique case, and finishing the feedback process in a timely fashion. The situation becomes even more complicated when the necessity of fairly assessing all solutions is added. In other words, although high quality feedback is as beneficial as a “breakfast of champions”, providing effective feedback in an open-ended problem-solving setting is not as easy as preparing breakfast.

To address the challenge, instructors should be able to quickly and accurately identify student misconceptions or other areas in student solutions that need support. Even with these skills, instructors may not be able to provide effective feedback. A typical example is providing feedback that is too long or complicated causing inattention in most learners. Thus, instructors should also be aware of how to communicate their feedback effectively. On the student side, students need to know how to interpret this feedback to get the maximum learning benefit and
produce improved products. Effective instructor-student interactions with feedback require well established, research-informed approaches that can be used to train and strengthen both sides of the feedback interaction. These approaches would ease the feedback process for open-ended problems and pave the way for multiple formative feedback steps in a single problem solving activity. The prerequisite for developing these approaches is an understanding of how students respond to feedback they are provided. The purpose of this study is to explore student team responses to teaching assistants’ (TAs’) written feedback while working on a mathematical modeling problem. The research question that guides this study is: How do student teams respond to feedback and convey their ideas from their team discussions in their documented works?

II. Background

Feedback can be defined as “information provided by an agent… regarding aspects of one’s performance or understanding”\(^3\) or as “information communicated to the learner that is intended to modify his or her thinking or behavior to improve learning”\(^2\). In the educational setting of this work, agents are instructors, specifically TAs, and the learners are first-year engineering students. TAs provide feedback on students’ drafts for the purpose of helping them narrow the gap between actual and reference level performance. A rubric and supporting materials provided to instructors before the mathematical problem solving activity explain the reference level performance.

There is an extensive body of literature on feedback which is summarized in a number of review articles\(^2\)–\(^6\). High quality feedback is generally regarded as an integral component of a fruitful learning experience especially when it is provided multiple times during the learning experience\(^4\). Feedback is reported as a significant factor in improving student knowledge, skills, and motivation\(^2\). Although much of this literature is not drawn from higher education settings where open-ended problem activities are implemented, the findings provide some generic guidance on the best practices for providing effective feedback. Shute offers a review of literature that includes a “to do” and “not to do” list for providing feedback\(^2\). In addition, she provides guidelines in relation to learner characteristics. For example, she mentions that low-achieving learners need directive (explicit guidance) feedback, whereas high-achieving learners need facilitative (hints, cues, prompts) feedback. In higher education, in spite of recognizing the importance of timely and effective feedback on student learning, feedback is not an area that has been widely addressed. Yorke emphasizes that “The importance of the student’s reception of feedback cannot be overstated.” and acknowledges the lack of literature on feedback\(^7\). Similarly, Mutch describes feedback to students as a “vital but relatively under-researched area” and calls for further research into how students receive and respond to feedback\(^8\).

Model-Eliciting Activities (MEAs) are realistic, open-ended, client-driven problems which are designed to develop students’ mathematical modeling skills along with communication and teamwork skills. These activities are constructed using six design principles: model construction, reality, self-assessment, model documentation, generalizability, and effective prototype\(^9\). These principles are modified for engineering contexts\(^10\). In these activities, students are expected to develop shareable, reusable, and modifiable mathematical models for clients. These activities are recognized as examples of effective engineering learning experience\(^11\).
In MEAs, feedback can be provided in a formative manner and multiple times during the course of problem solution development. This feedback can be provided by graduate TAs or peers. Yildirim et al. developed nearly 20 engineering-based MEAs and implemented them at various levels in different courses of a university. These include courses like “Decision Models” (student profile: mixed engineering disciplines/ graduate students and seniors), “Engineering Ethics” (student profile: mixed engineering disciplines/ juniors and seniors), “Open Ended Problem Solving” (student profile: industrial engineers/ juniors). They emphasize the importance of formative feedback: “We have learned that the success of the MEA depends on the guidance given students while working on the MEA.”

The research concerning the nature of feedback to student teams during MEA implementation is in its infancy. For example, research on written feedback on MEAs began with a framework developed by Diefes-Dux et al. in 2012. Despite the awareness of the influence of feedback on students’ success in mathematical modeling, there is little known about which components of feedback realize this influence. For example, in the broader literature, it is known that detailed feedback is beneficial. However, in mathematical modeling activities, the kind of detail in feedback that offers a better learning experience is still not well understood. The prerequisite for understanding the nature of effective feedback is to understand how students perceive and respond to various types of feedback they are provided. The feedback TAs and peers provide and the way students respond to it when revising their solutions capture important data about students’ thinking processes. In the literature, these thinking processes have been revealed mostly through an analysis of documented works, such as written feedback and student solutions. However, such approaches do not reveal the whole story of students’ interactions with feedback. The purpose of this study is to explore student teams’ responses to TAs’ written feedback while revising their mathematical model. As opposed to most engineering education MEA studies in the literature, real-time data of the student teams working with feedback to improve their models are included in this study. Such an approach is expected to reveal the stories behind students’ documented works, as exposed through their conversations while responding to feedback.

III. Method

A. Participants, Setting, and MEA Implementation

The data for this study was collected at Purdue University. Participants for this study were the 1,164 students enrolled in a required introductory problem-solving and engineering computer tools course. In this course, students, in teams of 3 or 4, solved MEAs, which are authentic engineering mathematical modeling problems.

All of 1,164 students completed three MEAs. This work focused on the second MEA of the semester: Just-In-Time Manufacturing (JITM) MEA. In this MEA, a fictitious manufacturing company (DDT) was unsatisfied with its current shipping service. Operating in a just-in-time manufacturing mode, the company was looking for a new shipping service to move materials between two locations. The company needed the student teams to develop a procedure to rank candidate shipping companies in order of best to least able to make deliveries on time. Student teams were provided historical data of the number of minutes late a shipment arrived at its destination. For the Draft 1 submission, students were provided a small subset of a larger
historical data set. For later stages, Draft 2 (focus of this work) and Team Final Response submissions, students were provided full historical data for 4 and 8 shipping companies, respectively. The full historical data for a shipping company included 270 arrival times in minutes late, ranging from 0 to 100 minutes. All eight shipping companies had varying data with similar means (9.90 min. +/- 0.4 min.). Therefore, students should have realized that the mean alone did not provide enough information to differentiate shipping companies. Moreover, no shipping company had the lowest mean and lowest standard deviation combination, which should have pushed student teams toward exploring other characteristics of the data set. Due to the characteristics of the data set, student teams’ solutions, formatted as memos, were expected to account for the distribution of these data as well as provide a reasonable strategy to break ties.

The semester started with a faculty lead introduction to MEAs and open-ended problem solving. This introduction aimed to show students that the goal of MEAs is not to solve the immediate problem (in the JITM case rank the eight shipping companies) but to develop a high quality, generalizable, and shareable procedure (model) that meets a client’s immediate and future needs. Solutions to each of the three MEAs were then developed according to the implementation sequence shown in Figure 1. First, individual students submitted an individual homework asking three free-response questions to identify the client, the need of the client, and issues to be considered when developing a solution. Second, student teams were provided a small subset of a larger historical data set in class, and they completed and submitted Draft 1 in no more than 90 minutes. Third, TAs provided student teams feedback on their Draft 1 memos using the MEA Feedback and Assessment Rubric. While a grade was indicated to the students at this stage, a final MEA grade was only issued at the end of the development process. At the same time, students attended a lecture on how to interpret feedback. After this, student teams had one week to revise and submit Draft 2 and Changes document (this prompts students to state what TA feedback they are addressing and how). Fourth, individual students completed online peer feedback training in which they were given one sample MEA solution to evaluate and then compare with an expert evaluation. Fifth, students individually were required to give feedback to one other peer team in a double-blind fashion. Sixth, student teams once again revised their work considering peer feedback they had received and submitted their Final Responses. Finally, JITM MEA sequence was completed with TA feedback on and assessment of the Team Final Response. TAs’ feedback on the Team Final Responses was intended to be helpful for successive MEA solution development. As shown in Figure 1, this work focused on the period in between student teams’ Draft 1 and 2 submissions.

TAs play an important role in MEA implementations. They often function as students’ primary contact, and they provide feedback to student work. Thus, TAs received basic training at the beginning of the semester and training about each MEA during the semester. Using face-to-face training, practice on prototypical student work and other methods, TAs were prepared to reliably evaluate and provide feedback on students’ work.
Figure 1 – *Just-In-Time Manufacturing* Model-Eliciting Activity implementation sequence.
B. Data Collection

This work used a case study research method\textsuperscript{21}. The purpose was to explore student team responses to TAs’ written feedback while revising their mathematical models.

The data collected in this study were student teams’ written works (Draft 1, Draft 2, and Changes document), TAs’ written feedback, video-recordings of teams working on revising their drafts, and semi-structured post interviews with students (Figure 2). Draft 1, Draft 2, the Changes document and TA feedback were all collected via an online MEA management system\textsuperscript{22}. For responding to the TA feedback and revising their Draft 1s, student teams were invited to a laboratory setting supported by laptops, internet access, and work space with whiteboards. Students’ teamwork, including their conversations and notes on whiteboards, were video-recorded.

Eleven teams volunteered to be videotaped as they responded to written feedback from TAs, and the individual team members committed to being interviewed following the MEA. After the collection and initial analysis of data, Team A and B were selected among the eleven teams. There were basically two rationales behind choosing these two particular teams: (1) they had nearly complete data sets (i.e., written documents, interviews, and video-recordings), and (2) both had similar TA feedback about lack of a method to address ties in their model and this
would have let us compare the way they respond to this specific feedback item. One more reason, which appeared after we picked these teams, was that the video-recordings cover their full Draft1 revision teamwork episode. Our concern was that the team would have read the TA feedback prior to the meeting and begun thinking about or even writing some portions of their Draft 2 before coming to the team meeting. Fortunately, detailed analysis of their video-recording revealed that both teams started working at the beginning of the video and submitted their Draft 2 at the end of the video.

As mentioned above, data focused on two student teams, Team A and B. For the purpose of this study, students were identified with a pseudonym as in Figure 3. Each team consisted of four members. Team A was an all-male team. Three out of four Team A students appeared in the video-recording. Student A-4 did not attend the team meeting. However, he did make a contribution - prior to the meeting he sent the team a small paragraph on addressing one of the feedback items (provided in Results section). Team B was a half male and half female team with all its members attending their team meeting.

After the submission of their Draft 2 and Team Final Responses, students were individually invited to an interview session. The interviews were semi-structured with guiding questions about their TA and peer feedback experience. Relevant guiding questions from the student interview protocol are provided in Table 1. Two Team A students and three Team B students completed the interview sessions. All sessions were audio-recorded.

Table 1
Relevant questions asked to students in semi-structured interviews
(From the interview protocol)

<table>
<thead>
<tr>
<th>Part 1: Experiences responding to feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can you please describe for me the type of feedback you received from your TA?</td>
</tr>
<tr>
<td>How did your team respond to this feedback? Would you have done anything differently if you were responding by yourself?</td>
</tr>
<tr>
<td>Did you encounter any challenges in responding to the feedback from your TA? If yes, how did you work around these challenges?</td>
</tr>
</tbody>
</table>
C. Data Analysis

Written documents for both teams set the stage for the focus of this case study. Comparison of student teams’ Draft 1 and Draft 2 submissions revealed what changes student teams made during their revision meeting. This was supported by their Changes document. TA feedback was parsed into separate items and similar items grouped under titles like: (addressing) TIE, (provide) CLEAR STATEMENT (detailed in Results section).

The feedback items identified in the TA feedback, were matched to the time frames of when students start and end discussing each feedback item in the video-recordings. The conversations in these frames were compared to the related part of their documented works to see how they conveyed their ideas. Portions of the video-recordings which were found to be related to the research question of this study were transcribed and used to triangulate findings from the written documents and interviews.

Interviews were transcribed and relevant pieces about how they respond to TA feedback were selected. These pieces provided insights to students’ individual reflection on their teamwork.

Results of Team A and Team B were constantly compared with each other to identify the similarities and differences in the ways they respond to TA feedback and conveyed their ideas in their documented works.

IV. Results

Submitted documents, TA written feedback, student interviews, and video data were analyzed to answer the research question of this study: How do student teams respond to feedback and convey their ideas from their team discussions in their documented works? Teams’ actual responses to TA feedback were identified from their submitted documents (Draft 2 and Changes documents) and video-recording of their meeting. Interviews revealed not only their perception of TA feedback but also their reflection on their response.

A. Team A Case

Team A had a simple mathematical model in their Draft 1. For each shipping company, they calculated the mean (average) of the provided historical data of the number of minutes late a shipment arrived at its destination. They also identified maximum and minimum delivery times. Then, they multiplied Mean by Max – Min. They provided no rationale for multiplication. Their mathematical model, Mean × (Max – Min), poorly accounted for data variability and did not attend to the distribution of the data. Their Draft 1 had other issues like not providing a method to resolve ties. These issues as well as others were highlighted in the TA’s written feedback (Table 2). Table 2 lists the codes used in this study for each feedback item. These codes are used in the following tables and figures.
Table 2  
TA written feedback items to Team A

<table>
<thead>
<tr>
<th>Feedback items</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>- You did a really good job at providing rationales for your steps</td>
<td>Not coded</td>
</tr>
<tr>
<td>- How does your procedure process missing data?</td>
<td>MISSING DATA</td>
</tr>
<tr>
<td>- You need to provide a method to resolve ties.</td>
<td>TIE</td>
</tr>
<tr>
<td>- provide tie breakers</td>
<td></td>
</tr>
<tr>
<td>- Are there any other statistical measurements that might provide your ranking procedure with more information about the distribution of the data?</td>
<td>DISTRIBUTION</td>
</tr>
<tr>
<td>- Are there any cases that you can think of that might challenge your procedure? Explain the limitations of your procedure and provide assumptions accordingly.</td>
<td>LIMITATIONS</td>
</tr>
<tr>
<td>- Any other assumptions explicit or implicit in your procedure?</td>
<td></td>
</tr>
</tbody>
</table>

The TA also evaluated Draft 1 according to *MEA Feedback and Assessment Rubric* \(^{17}\). There were seven grades associated with different aspects of Team A Draft 1, for the purpose of this study, it is important to note that Team A would have received a C grade based on the quality of their mathematical model.

Students perceived some aspects of the TA feedback differently (Table 3). Although their draft was mathematically problematic, Student A-3 found the TA feedback positive and thought that they did not have much to change. He said, “*We made a few changes. We made it a little more complex in our solution just finding like, the ranking, finding out how to write the different shipping companies. But other than that… this time we had a lot better rationales for everything so [TA] said we did perfect on that, that we didn't need to change anything with that so we were pretty well off after the first draft so didn’t have much to change.*”

Table 3  
How Team A perceived TA feedback

<table>
<thead>
<tr>
<th>Student A-1</th>
<th>Student A-3</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>It was kind of generic</em></td>
<td><em>It didn’t seem like it was a generic response.</em></td>
</tr>
<tr>
<td><em>We’re never sure what they want… that was a challenge</em></td>
<td>*It was a lot clearer this time… and a lot more</td>
</tr>
<tr>
<td></td>
<td><em>direct to our solution</em></td>
</tr>
</tbody>
</table>

Team A used the TA feedback to guide their revision meeting. Student A-1 explained how they responded to TA feedback as “*We just did what [TA] said or … how we interpreted it so that's how we respond to TA feedback.*” The video analysis revealed that they started with the
MISSING DATA feedback item. This was because Student A-4, who was absent from the meeting, had sent them a paragraph on addressing the MISSING DATA issue. He wrote, “If data is missing, it’s assumed the time was not recorded. If a product was not delivered, assume that the data would say "not delivered." If a product is never delivered, the company cannot be chosen as a deliverer. If there is missing data, do not take any account for it as it is not the fault of the delivering companies.”

This small paragraph was the only contribution of Student A-4. At the beginning of the meeting, Team A read this email and then, decided to copy this explanation into their draft. Thereby, they finished addressing MISSING DATA feedback in less than 3 minutes. After this, they moved to the TIE feedback and again very quickly (6 minutes) devised a solution: using (Max-Min) to break ties. Their thinking process throughout the meeting was not deep, which was highlighted by Student A-1 as “We’re like: “ah, sounds good”, “Let’s just go with that” and that’s about how that (our team) works.” However, when they moved to the DISTRIBUTION feedback, their pace slowed and a lengthy discussion of DISTRIBUTION started. This discussion was interrupted with discussion of other feedback items; however, addressing distribution of the data remained at the core of their meeting. Video analysis revealed that they spent 52 minutes in total addressing DISTRIBUTION. Details of this analysis are presented in Table 4.

Table 4
Team A meeting progress

<table>
<thead>
<tr>
<th>Start time hr:min:sec</th>
<th>End time hr:min:sec</th>
<th>Feedback Item</th>
<th>Minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:06:10</td>
<td>00:07:27</td>
<td>MISSING DATA</td>
<td>1.28</td>
</tr>
<tr>
<td>00:08:00</td>
<td>00:08:38</td>
<td>MISSING DATA</td>
<td>0.63</td>
</tr>
<tr>
<td>00:08:58</td>
<td>00:15:06</td>
<td>TIE</td>
<td>6.13</td>
</tr>
<tr>
<td>00:17:09</td>
<td>00:19:40</td>
<td>DISTRIBUTION</td>
<td>2.52</td>
</tr>
<tr>
<td>00:20:57</td>
<td>00:53:07</td>
<td>DISTRIBUTION</td>
<td>32.17</td>
</tr>
<tr>
<td>00:55:13</td>
<td>01:00:38</td>
<td>DISTRIBUTION</td>
<td>5.42</td>
</tr>
<tr>
<td>01:00:38</td>
<td>01:02:21</td>
<td>TIE</td>
<td>1.72</td>
</tr>
<tr>
<td>01:02:21</td>
<td>01:03:48</td>
<td>DISTRIBUTION</td>
<td>1.45</td>
</tr>
<tr>
<td>01:04:40</td>
<td>01:05:33</td>
<td>LIMITATIONS</td>
<td>0.88</td>
</tr>
<tr>
<td>01:05:33</td>
<td>01:06:46</td>
<td>DISTRIBUTION</td>
<td>1.22</td>
</tr>
<tr>
<td>01:06:46</td>
<td>01:07:16</td>
<td>LIMITATIONS</td>
<td>0.50</td>
</tr>
<tr>
<td>01:07:20</td>
<td>01:16:53</td>
<td>DISTRIBUTION</td>
<td>9.55</td>
</tr>
<tr>
<td>01:35:18</td>
<td>01:36:39</td>
<td>TIE</td>
<td>1.35</td>
</tr>
<tr>
<td>01:37:01</td>
<td>01:50:10</td>
<td>LIMITATIONS</td>
<td>13.15</td>
</tr>
</tbody>
</table>

Team A’s progress is also illustrated in Figure 4. In this figure, total time spent visiting each feedback item is shown near the associated code boxes.
As seen in Figure 4, DISTRIBUTION was the feedback item which took the most of Team A’s meeting time. The video analysis revealed that the outputs of their DISTRIBUTION discussion feed the discussions of two other feedback items, TIE and LIMITATIONS. The prominence and centrality of the DISTRIBUTION discussion is seen in Figure 4 where discussion bounces back and forth between DISTRIBUTION, TIE and LIMITATIONS. Thus, it is important to identify what happened during this discussion.

TAs were instructed to guide students towards a mathematical approach rather than tell them what math to use. This was done in the DISTRIBUTION feedback item provided to Team A: “Are there any other statistical measurements that might provide your ranking procedure with more information about the distribution of the data?”

This feedback was sufficiently effective for guiding Team A to improve their math model. However, their video revealed that they decided to change their model not because their earlier model needed change. Rather, their motivation was just doing what the TA had said and somehow getting a better grade. Some quotes from their initial conversations are as follows:

A-1: you know they teach us all that nonsense like standard deviation, mean, median, average, mode as?
A-2: Yeah
A-1: Do you want to try and apply some of them?
A-3: I am thinking about it, like a bell curve or something like that

...  
A-1: Let’s think of other stuff, that is the standard deviation. I am gonna go get my notes. Because the more stuff we input from class...
A-3: The more gonna they will be freaking like “awesome, good”
(A-1 opens his course notes)
A-3: well, just draw a best fit line for the data or something, they will love that
A-1: dude, if we draw a best fit line for this data, we will get automatic million percent.

After some time, they decided to use variance (or standard deviation), but the team was unsure of what this measures.

A-1: I don’t even know what this variance gives us, just sounds good. ... As long as we are like “well, we can also determine variance” and put that or something like that, they will be like [happy].

... 

A-3: why do we include standard deviation?
A-2: that’s what I’m trying to say
A-1: better judge of consistency
A-2: what we have to say is that if you take only two standard deviation out, that way it doesn’t include the outliers...
A-3: that’s a good idea
A-1: what’s the definition of standard deviation?

After these conversations, they decided to change their model from Mean × (Max – Min) to Mean × (Max – Min) × 2 Std. dev. They added the following explanation to their Changes document:

We're including standard deviations in the mathematical model by multiplying the consistency data with the standard deviation of each data set per company. We’re including standard deviation because it will be a good judge of consistency in relation to average.

They did not provide a clear definition of “consistency” in the Changes document or Draft 1. Depending on the context, they use it for (Max-Min) or Std. dev.. The explanation they provided in Draft 1 included more vague terms like “minimizing the significance of” or “skew the data”:

... Step Four: Multiply the number obtained from step three by two standard deviations (96% of the data is included in this). This method will give you a more accurate measurement of the companies shipping time consistency because it will minimize the significance of the maximum minus minimum values. The max-min value could skew the data even if there’s only one large outlier, but this value still should be taken into consideration.

The video data revealed that most of their DISTRIBUTION discussion was about understanding what standard deviation, the term which “sounds good”, does. While they were discussing distribution, they suddenly came up with an idea about addressing ties:

A-1: Oh, let’s use standard deviation as consistency... they will be like [happy].

Here “consistency” refers to their previous model for breaking a tie, which was (Max-Min). They do not talk much about this idea and move back to discussing DISTRIBUTION. Later in the meeting, while trying to figure out what standard deviation does, they revisit TIE:

A-2: we need to tell them (TAs) what it (standard deviation) does
A-1: (looking at computer screen) … “a measure of dispersion” … “standard deviation is a measure of how spread out your data is”

A-3: so the more consistent data, right?

A-1: the lower the standard deviation, the more consistent the data... we can do tie breaker by standard consistency then

A-2: by standard deviation?

A-1: yeah...

A-2: so are we just doing that for the tie breaker?

A-1: yeah. we can just use standard deviation for tie breaker because it is really hard to have tied standard deviations

A-2: yeah

Then, they agreed on using standard deviation to break a tie. They reported that, “We do this because consistency should be valued more”. They thought that if a shipping company is consistently X minutes late, then the shipment can be sent X minutes earlier than usual so that it arrives on time.

Overall, their Draft 2 still did not explain why they used multiplication in their math model (why not addition, for example). They claimed 96% of the data was included within two standard deviations. This was a misconception as (1) this assumes that the data is distributed normally (2) even under normality assumption, 96% of the data should be included in “mean + two standard deviations”, not in two standard deviations.

B. Team B Case

When compared to Team A, Team B had a more sophisticated math model in their Draft 1. They defined a parameter called “Std. dev. from zero”, which was standard deviation calculation as if the mean was zero. They also calculated the Median of the provided historical data. Then, they took the average of “Std. dev. from zero” and Median. Their mathematical model, \( \frac{1}{2} \times (\text{Std. dev. from zero} + \text{Median}) \), was a fair attempt at accounting for distribution of the data. Therefore, they did not get TA feedback which required them to change core of their mathematical model. However, they received feedback requesting that they provide a method to resolve ties and better explain their model. Student B-2 explained this as, “[TA Feedback] really, like, changed what we did in our, not so much as our procedure, but the way we, like, drafted our memo was completely different.” According to MEA Feedback and Assessment Rubric \(^{17}\), Team B would have received a B grade based on the quality of their mathematical model.

Table 5 lists the TA feedback items Team B was provided and how these items were coded in this study. Some items are “Not coded” since they were addressed easily, without much discussion during the team revision meeting. The codes in Table 5 are used in the following tables and figures.
Table 5
TA written feedback items to Team B

<table>
<thead>
<tr>
<th>Feedback items</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Your procedure does not address ties.</td>
<td>TIE</td>
</tr>
<tr>
<td>- Could this procedure be use by DDT if the shipping and receiving points were farther apart, say 120 miles or 300 miles or 800 miles. How far apart are Noblesville and Delphi, IN?</td>
<td>LIMITATIONS</td>
</tr>
<tr>
<td>- A clear statement describing the basis for the ranking system needs to be provided before the procedural steps.</td>
<td>CLEAR STATEMENT</td>
</tr>
<tr>
<td>- Better explain &quot;standard deviation from zero- penalizes outliers&quot;</td>
<td>BETTER EXPLAIN</td>
</tr>
<tr>
<td>- Your procedure does not clearly articulate how to evaluate &quot;standard deviation from zero&quot;, provide some calculations that will better articulate your procedure.</td>
<td>PROVIDE SAMPLE CALC</td>
</tr>
<tr>
<td>- Do not include &quot;procedure of how to import data into Excel and working with excel&quot; and &quot;procedure to calculate basic statistical measures&quot;, these are extraneous information.</td>
<td>Not coded</td>
</tr>
<tr>
<td>- Summarize only the results (company rankings with &quot;calculated sums&quot;).</td>
<td>Not coded</td>
</tr>
</tbody>
</table>

Team B perceived TA feedback positively and reported that they were “all happy with it”. Their perception as reported in their interviews is summarized in Table 6.

Table 6
How team B perceived TA feedback

<table>
<thead>
<tr>
<th>Student B-1</th>
<th>Student B-2</th>
<th>Student B-4</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>The errors [TA] made us address were a lot more minor</em></td>
<td><em>[TA] actually gave us, like, specific things that we needed to think about</em></td>
<td><em>This time it was a lot more positive than last time</em></td>
</tr>
<tr>
<td><em>There wasn't anything I thought was ridiculous</em></td>
<td><em>We were all happy with it. It guided us in the right direction. So, it was good.</em></td>
<td><em>We thought it was good. We agreed with every single point he had</em></td>
</tr>
<tr>
<td><em>It was straightforward, just correct this, correct that and get outta there.</em></td>
<td><em>There were open ended things, too...</em></td>
<td><em>[TA] was trying to get down the basics</em></td>
</tr>
</tbody>
</table>
Team B responded to their TA’s feedback with an intention to carefully address every part of it. Student B-1 explained their intention as “we wanted to make it obvious that we actually worked hard”. Their team meeting resulted in an intense discussion on various aspects of the MEA. Although they were interviewed separately after the meeting, each student agreed in interviews that they were involved in a deep discussion.

Table 7
How Team B responded to TA feedback

<table>
<thead>
<tr>
<th>Student B-1</th>
<th>Student B-2</th>
<th>Student B-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>We did kind of just an overall more serious and we tried to think about the problem a little bit deeper.</td>
<td>We actually had to, kind of, think more in depth into it.</td>
<td>There was a lot more brainstorming involved, a lot more arguing, a lot more people playing devil’s advocate. So we definitely went more in depth…</td>
</tr>
</tbody>
</table>

The video-recording of their meeting revealed that they started by writing feedback items on the whiteboard and trying to address each item “in depth” (B-2) and by “playing devil’s advocate” (B-4). Also, when they could not respond to an item, they left it and moved to another. This made the progress of their meeting very sophisticated as seen in Figure 5 and Table 8. In Figure 5, total time spent visiting each feedback item is shown near the associated code boxes.

Figure 5 – Team B’s meeting progress
As seen in Table 8, Team B started and ended their conversation with a discussion of the TIE feedback. Throughout their meeting, they discussed TIE eight different times (Table 8), and they spent the longest time on this feedback item (Figure 5). Thus, it is important to understand why Team B had to visit TIE many times and spent a significant amount of time on it.

The video revealed that interaction between the team’s discussion of the TIE feedback and the other feedback was limited. The reason why they visited TIE so many times was not because they applied the output of their TIE discussion to other feedback items and came back to TIE again. Rather, in each TIE visit, they failed to find a reasonable solution and they postponed their discussion to a later time.

**B-3:** We’re kind of avoiding ties.
**B-2:** I know. We were sort of talking about them and—
**B-4:** Do you want me to throw in average and see what happens?
**B-1:** Is there anything else that we can use that is worthwhile?
**B-4:** The mode (laughing by all).
**B-2:** I don’t like this whole tie thing.
**B-1:** I don’t either.

<table>
<thead>
<tr>
<th>Start time hr:min:sec</th>
<th>End time hr:min:sec</th>
<th>Feedback Item</th>
<th>Minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:04:10</td>
<td>00:04:16</td>
<td>TIE</td>
<td>0.10</td>
</tr>
<tr>
<td>00:11:04</td>
<td>00:11:45</td>
<td>TIE</td>
<td>0.68</td>
</tr>
<tr>
<td>00:13:21</td>
<td>00:17:19</td>
<td>LIMITATIONS</td>
<td>3.97</td>
</tr>
<tr>
<td>00:17:55</td>
<td>00:18:35</td>
<td>LIMITATIONS</td>
<td>0.67</td>
</tr>
<tr>
<td>00:22:19</td>
<td>00:22:33</td>
<td>TIE</td>
<td>0.23</td>
</tr>
<tr>
<td>00:34:48</td>
<td>00:37:32</td>
<td>TIE</td>
<td>2.73</td>
</tr>
<tr>
<td>00:37:32</td>
<td>00:44:23</td>
<td>LIMITATIONS</td>
<td>6.85</td>
</tr>
<tr>
<td>00:44:24</td>
<td>00:46:50</td>
<td>CLEAR STATEMENT</td>
<td>0.43</td>
</tr>
<tr>
<td>00:44:40</td>
<td>00:46:32</td>
<td>LIMITATIONS</td>
<td>1.87</td>
</tr>
<tr>
<td>00:46:32</td>
<td>00:47:13</td>
<td>BETTER EXPLAIN</td>
<td>0.68</td>
</tr>
<tr>
<td>00:47:13</td>
<td>00:47:30</td>
<td>PROVIDE SAMPLE CALC</td>
<td>0.28</td>
</tr>
<tr>
<td>00:47:31</td>
<td>00:48:43</td>
<td>LIMITATIONS</td>
<td>1.20</td>
</tr>
<tr>
<td>00:48:44</td>
<td>00:49:45</td>
<td>TIES</td>
<td>1.02</td>
</tr>
<tr>
<td>00:49:45</td>
<td>00:50:35</td>
<td>CLEAR STATEMENT</td>
<td>0.83</td>
</tr>
<tr>
<td>00:54:04</td>
<td>00:54:17</td>
<td>TIES</td>
<td>0.22</td>
</tr>
<tr>
<td>00:55:07</td>
<td>00:56:08</td>
<td>LIMITATIONS</td>
<td>1.02</td>
</tr>
<tr>
<td>00:59:28</td>
<td>01:09:48</td>
<td>BETTER EXPLAIN</td>
<td>10.33</td>
</tr>
<tr>
<td>01:09:50</td>
<td>01:10:43</td>
<td>CLEAR STATEMENT</td>
<td>0.88</td>
</tr>
<tr>
<td>01:12:08</td>
<td>01:16:10</td>
<td>TIE</td>
<td>4.03</td>
</tr>
<tr>
<td>01:16:35</td>
<td>01:27:17</td>
<td>PROVIDE SAMPLE CALC</td>
<td>10.70</td>
</tr>
<tr>
<td>01:27:18</td>
<td>01:40:43</td>
<td>TIE</td>
<td>13.42</td>
</tr>
</tbody>
</table>
But we have to address it
We have to address it or our mathematical model—
—is going to fail.

As seen in their discussion, they always looked for additional measures (Mode, Mean, etc.) to incorporate into their existing model; whereas Team A did the opposite; they used Mean, Max, Min, and Std. dev in their math model and for the tie breaker, they just used Std. dev. Student B-4 explained their experience as “We thought we already used all the mathematical parts we could on our normal procedure, so we didn’t know how to break a tie...we found that (tie) was almost impossible unless they had the exact same data, to have it tied, the way we did it.”

The video revealed that in their eighth and last TIE discussion, at about 87 minutes into their meeting, they discussed incorporating Average (Mean) into their tie breaker model in addition to “Std. dev. from zero” and Median. In such a case, the model they used for tie was going to be $1/3 \times (\text{“Std. dev. from zero”} + \text{Median} + \text{Mean})$. They changed the constant from 1/2 to 1/3 because this time, they wanted to take average of the three parameters. After discussing this tie breaker model, they gave up since (1) they could not find a reason for using Mean in addition to Median and (2) they realized that it was very difficult to get a tie with their existing math model:

Wait. We need to have a rationale for why we decided to just add the average in.
I was about to say we have to have a rationale for the average which is going to be difficult.
We’re just like, let’s throw that in there and see what happens and then we have to figure out the rationale. We’re like working backwards.
That’s okay though. We’ll see if we like the results.
We’re doing what engineers do a lot of the time. You give them something that has to accomplish something, but they don’t know what that things is and what it looks like at all. They’re just given a goal and you have to figure out what accomplishes that goal. We had to make a tie breaker and we had to figure out something that actually breaks ties. Although we don’t have ties to begin with, but whatever.

It’s hard to tell because we don’t actually have ties to break.

Finally, they discussed a simpler solution: the logistic manager will use personal preference. The video analysis revealed that they didn’t like this idea, but could not agree on any better solution since they realized that their mathematical model avoided creating ties in the first place

But the only way to produce a tie is to have something that has identical data. At that point you can say it becomes the personal preference of the company.
It just makes me really nervous is the personal preference thing.
It seems like almost impossible to have a standard deviation the same and the median.
Even though some are really close, they’re not identical. They’re at least a thousandths off.
They’re even more than that—at least a tenths off of each other.
B-2: Which could be some money. So that got us somewhere.

... 
B-1: We can say that the only way that tie breaker would be [required] would be if you were using nearly identical data points and that in case it does come down to personal preference.
B-2: We have no other data to work with. They don’t give us the cost or anything about the individual companies.
B-4: So what we’re saying is that we can’t come up with a more accurate mathematical model than this, so you’re going to have to deal with it.
B-1: No, we’re saying that using our mathematical model, ties just don’t really happen. Not that ours is the best or the most accurate. It’s just the fact that the sort of calculations that we use. We make this calculation look too simple. It’s not just two numbers added together.

The video analysis also revealed that they tried to simulate a tie scenario by producing their own data for two companies. However, they realized this was very difficult with a large data set and agreed that this would only be possible with almost identical data for two companies. In her interview, Student B-2 referred to this discussion:

... So we say if they had the exact same standard deviation and median, which was what we used, then (tie breaker is needed and) they probably just have to do it on preference. Which I’m not sure of that’s a good thing that we said that or not. Probably not, but we couldn’t think (laughter) of anything.

Thus, instead of using a separate step to eliminate ties if they happened, they decided to add a modest explanation in their memo:

On a side note, we need to address the situation of a tie occurring. The only way a tie could be produced using our mathematical procedure would be to use nearly identical data sets (due to the use of standard deviation from zero). Therefore, in the rare case of a tie, the logistics manager may have to consider factors such as cost and overall personal preference.

Before concluding this part, a summary of both teams’ solutions is provided in Table 9.

Table 9
Team A and B solutions

<table>
<thead>
<tr>
<th>Team A Draft 1</th>
<th>Team A Draft 2</th>
<th>Team B Draft 1</th>
<th>Team B Draft 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean × (Max – Min)</td>
<td>Mean × (Max – Min) × 2 Std. dev.</td>
<td>½ × (“Std. dev. from zero” + Median)</td>
<td>½ × (“Std. dev. from zero” + Median)</td>
</tr>
<tr>
<td>-</td>
<td>Std. dev.</td>
<td>-</td>
<td>External factors (cost, personal preference etc.)</td>
</tr>
</tbody>
</table>

V. Discussion
This study focused on how student teams responded to TA written feedback and conveyed their team discussions in their documented works. While this was a case study concerning two student teams, the results can provide insights into the broad literature on the complexity of engineering students’ learning experience.

Team A’s response to their TA’s feedback showed that they were very focused on making the TA happy, and ultimately getting a good grade. This lead to blindly employing statistical measures they saw in class without much attempt to better understand the concepts and how they might be employed to solve this problem. Their resulting model (equation) and rationales indicated their insufficient understanding of data distribution. The team did not realize that standard deviation does not provide sufficient information about how data is distributed for the context of the given problem. The lecture discussion on the bell shape of a normally distributed data set and how standard deviation helps describe that shape were not understood. This resulted in the comment about 96% of the data being included within two standard deviations. This team did not test the data sets they were provided for normality.

Hattie and Timberley\(^3\) discuss that the way students receive feedback is dependent on their characteristics (e.g. self-efficacy). Students construct their own meaning out of the provided feedback. This was observed in Team A’s case (Table 3). Student A-1 said that the feedback their team received was “kind of generic”; whereas Student A-3 said it did not seem generic. In addition, they disagreed about the clarity of the feedback. Thus, as stated by Shute\(^2\), it is important to take individual characteristics of learners into account when composing a feedback message.

However, individual characteristics of students may not directly appear in students’ written document. Nevertheless, Team A’s TA could have a rough idea concerning their performance level by looking at their weak rationales and ambiguous use of terms like “consistency”, “minimizing the significance of”, and “skew the data”. Accurate and early identification of low performance is especially important because, as Shute\(^2\) mentions in her literature review, low-achieving learners need directive (explicit guidance) feedback. If TAs can correctly identify low-achieving learners while assessing their Draft 1s, they could write more explicit feedback.

For example, Team A received an implicit feedback comment about DISTRIBUTION (“Are there any other statistical measurements that might provide your ranking procedure with more information about the distribution of the data?”). As a response to this implicit feedback, the team opened their class notes and looked for something that “sounds good” without questioning why they got such a feedback or what was missing in their existing model.

A-1:  ... I am gonna go get my notes. Because the more stuff we input from class...
A-3:  The more gonna they will be freaking like “awesome, good”

Team A spent about 52 minutes on this implicit feedback, and at the end they still did not have a ranking procedure which adequately addresses data distribution. Team A’s learning experience might have been more efficient and productive if the TA had provided more explicit feedback. The TA could have mentioned that central tendency and variation (i.e. range) are accounted for in Team A’s Draft 1 math model, but not distribution. Moreover, in very low-achieving cases
like Team A, TAs might clearly explain that standard deviation is a measure of variation but not distribution, as experience with the JITM MEA tells us that students are often fixated on mean and standard deviation\textsuperscript{16,20}.

This analysis showed that TAs need to better identify misconceptions and target feedback appropriately. Another issue about the TA’s written feedback was the lack of warning about units. Team A Draft 1 model was an equation the results of which were not interpretable. The unit of the output was \textit{min}\textsuperscript{2}. It is likely that the students were trying to get the two measurements (\textit{Mean} and \textit{Range}) to play off each other by multiplying them – the product of a large mean and large range would be less desirable than a small mean and small range. However, combinations of large and small means and ranges would be difficult to interpret. The TA could have warned the students that multiplication of the measures concerned was not appropriate method to rank the companies. This warning could be supported by an example showing how the math employed by students was not meaningfully addressing the problem.

When compared with Team A, Team B was comprised of higher achieving learners. An indication of this was, while Team A did not know the definition of standard deviation at the beginning of their meeting, Team B came to the meeting with a modified standard deviation called “standard deviation from zero” in their Draft 1. Modifying the standard deviation to suit the needs of the problem at hand can be considered a fairly good conceptual understanding of what standard deviation does.

Team B really tried to make sense of the TA feedback. Student B-1 explained their response as “we wanted to make it obvious that we actually worked hard”. While both teams got similar explicit feedback about TIE, which requested a tie breaker, Team B did not only look for a tie breaker but also spent considerable time trying to think through how the data might even produce a tie. As seen in their video recording, they could not understand why they needed a tie breaker since their model hardly resulted in a tie. As explained in the literature, high-achieving learners need facilitative (hints, cues, prompts) feedback rather than directive (explicit guidance) feedback\textsuperscript{2}. The TIE feedback could be classified as directive feedback since it explicitly guided students to provide a tie breaker. This explicit feedback, in a sense misguided Team B, since they realized that their math model does not need a tie breaker and had difficulty in understanding why they needed to provide a tie breaker. Team B might have responded much less painfully, if they had a facilitative feedback like “Do you think that your math model needs a tie breaker?”. However, this should not mean that high-achieving learners should be provided all-facilitative feedback. Team B also had explicit feedback like PROVIDE SAMPLE CALC feedback and easily responded to it.

Video analysis of the tie discussion in Team B’s meeting uncovered a problem which would not be seen by just examining students’ work: Team B students could not convincingly convey their thinking on TIE in their documented works. Thus, the evidence of their thinking was not apparent in their work to be evaluated by TAs. This problem was discussed in Team B’s meeting:

\textit{B-1: We’re already doing better than we did last time.}\n
\textit{B-2: That’s good. We’re showing improvement.}
B-1: I wish we could get credit for this.
B-2: They [TAs and profs] don’t see this.
B-1: That [we] are sitting here thinking about this critically.

Thus, it is very important to provide more instruction to students on how to show evidence of their thinking and work. Perhaps, they could have used the Changes document to explain their thinking processes not expressing in their memo. For example, they could have written “We discussed adding Mean to our current model by using $1/3 \times (\text{Std. dev. from zero} + \text{Median} + \text{Mean})$; however, we thought we already addressed the central tendency with using Median in our current model. This is why we keep our model the same in Draft 2.” In addition, Team B could have used the Changes document to explain how they tried to achieve a tie with their model and failed because it was almost impossible to get a tie in their case. They also could have provided example results in the Changes document to justify their assertions.

Moreover, Team B could have better interpreted the TIE feedback and budgeted their time accordingly. They might have realized that a tie breaker was not the central issue of this MEA and would not have spent most of their time on that. Actually, although their model was close to addressing the needs of the problem by using Median (a measure central tendency) and “Std. dev. from zero” (a fair measure of distribution), it would have been better if they could have found a more interpretable way of employing them. Averaging these two measures did not produce a measure that makes sense; rather, it seemed to indicate that the team believed the two measures were measuring the same thing (like taking the average of two measures of central tendency). It would have been much better if they had taken this issue to be the central focus of their meeting. Overall, Team B’s case implied that students need to learn how to interpret feedback, budget time, and effectively report the outcomes of team discussions.

Furthermore, Team B case revealed that TAs should not strictly look for a tie-breaker. They need license, through their grading guide and training, to examine arguments like the “tie is rare in our model” for validity. It might be possible to add a remark in the TA guide to allow for reasonable explanations that a tie is not likely to happen. In such cases, not providing a tie breaker would be considered acceptable.

VI. Conclusion

The video data revealed several important features of students’ interactions with TA’s written feedback, which could not be seen by examining students’ written work alone. To illustrate, time spent on each feedback item, how students come to the math model they report, the complexity of their discussions, the level of their knowledge, and how their discussion on a feedback item interacted with their discussions of the other feedback items could be uncovered thanks to video data. This study implies that video data based research results, which cannot be obtained by studies depending only on the written work, could lead to new research-informed 1) instructor training approaches to providing feedback, and 2) student instruction on how to interpret and respond to feedback.

This case study addressed two student teams’ interactions with TA feedback. Team A had a low quality model compared to Team B. Team A was quite grade-focused and was thus not very
engaged in learning through the open-ended problem and associated feedback. On the other hand, Team B was quite eager to make it obvious that they worked hard. Thus, the teams’ interactions with feedback were pretty disparate. Future studies could look at additional teams that perhaps spanned the quality continuum between these two teams to better capture how teams work with feedback and whether or not different strategies help or hinder progress towards an improved model or understanding of the embedded concepts.

Some evidence that misconceptions exist about statistical measures, like standard deviation, were obtained. If further research with additional teams finds that these results are generalizable, the implications for addressing these misconceptions in lectures could be discussed.

High quality feedback is fundamental in open-ended problem implementations to increase efficiency and effectiveness of students’ learning process. Since feedback should be specific to the achievement level of the learner, instructors need to be able to think through what the math the students employ really means and what misconceptions they have. Then, they need to target their feedback appropriately and employ effective strategies for getting students to better understand the concepts. However, students also need to know how to respond to the feedback and convey their thinking processes through their work. The research on feedback concerning all of these dimensions still merits further exploration to realize feedback’s powerful potential to influence learning and achievement.

Acknowledgement

This work was made possible by grants from the National Science Foundation (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


