Enhancing TA Grading of Technical Writing: A Look Back to Better Understand the Future

Dr. Krista M. Kecskemety, Ohio State University

Krista Kecskemety is a Senior Lecturer in the Engineering Education Innovation Center at The Ohio State University. Krista received her B.S. in Aerospace Engineering at The Ohio State University in 2006 and received her M.S. from Ohio State in 2007. In 2012, Krista completed her Ph.D. in Aerospace Engineering at Ohio State. Her engineering education research interests include investigating first-year engineering student experiences, faculty experiences, and the connection between the two.

Mr. Andrew H. Theiss, Ohio State University

Andrew Theiss is a Ph.D. student in the biomedical engineering graduate program at The Ohio State University. Andrew received his B.S. in Electrical Engineering at The Ohio State University in 2009. He currently works as a graduate research associate in the Wexner Medical Center and is in his third year as a graduate teaching associate in the Engineering Education Innovation Center (EEIC). His engineering research interests are focused on the development of biosensing systems utilizing field-effect transistor platform, and his education research interests include first-year experience and TA development.

Dr. Rachel Louis Kajfez, Ohio State University

Dr. Rachel Louis Kajfez is an Assistant Professor of Practice in the Engineering Education Innovation Center and the Department of Civil, Environmental, and Geodetic Engineering at The Ohio State University. She earned her B.S. and M.S. degrees in Civil Engineering from Ohio State and earned her Ph.D. in Engineering Education from Virginia Tech. Her research interests focus on the intersection between motivation and identity of undergraduate and graduate students, first-year engineering programs, mixed methods research, and innovative approaches to teaching. Currently, she teaches within the first-year engineering program at Ohio State while maintaining an active engineering education research program.
Enhancing TA Grading of Technical Writing: 
A Look Back to Better Understand the Future

Abstract
Technical writing is an important skill in all engineering disciplines. Many first-year engineering programs (FYEPs) include technical writing as one of their core course components to begin to instill the importance of and to develop this skill early in aspiring engineers. In order to assess student learning and provide feedback on technical writing, proper grading of these assignments is essential. This paper presents the preliminary assessment results of a new grading training program for teaching assistants (TAs) in a FYEP that was implemented at a large land grant institution while situating the findings in past research related to the grading of writing assignments and preparing TAs. This paper summarizes the approach used to enhance the grading of TAs. We supplement this discussion with quantitative and qualitative assessment results. Finally, we provide recommendations or “Tricks of the Trade” for TAs who may be grading technical writing and those who are responsible for training TAs for this type of grading.

Introduction
Engineering programs throughout the country are increasing the amount of technical writing in their curriculum because technical communication is recognized as a highly valued skill for engineers of all disciplines. While collaboration with both university and departmental writing centers can be valuable, the inclusion of technical writing in the context of engineering classes is essential. As the amount of technical writing increases in the classroom, it is important to train those responsible in providing feedback, whether it is faculty or TAs. Despite the recognized need, lack of training in addition to TAs or faculty feeling inadequate to provide constructive feedback can be a barrier to including effective technical writing in the classroom.

Generally speaking, graduate teaching assistant (GTA) training is important, as they are expected to perform many roles in the classroom, and they have many responsibilities including providing feedback and evaluating student work. However, there is no consistent method for training GTAs between universities and only 17% of universities spend more than one day on formal training activities. Broadly speaking, TA training efforts should be expanded to undergraduate teaching assistants (UTAs), should be beyond one day, and should include opportunities for practicing their teaching skills, which could include lecturing, grading, providing feedback, and classroom management.

At our large, land-grant university, the FYEP conducts courses for over 2000 students each year. All engineering students are required to take one of the two tracks (Fundamental of Engineering (FE) or Fundamentals of Engineering for Honors (FEH)). Each track teaches students problem solving, computer programming, technical graphics (visualization and sketching), CAD, and design. In addition to these topics, students are also exposed to many hands-on labs and taught fundamentals in technical writing and communication. In order to grade the technical writing assignments, approximately 50 GTAs and 150 UTAs are employed in the program each year. The grading of lab reports, memos, executive summaries, abstracts, and design reports is a difficult, time consuming task. Additionally, consistency between sections of a course is something that is of utmost importance. Because these course sections are intended to be as identical as possible, the consistency with grading (both in scores and feedback) is crucial for the
program to succeed. Therefore, there is a need for substantial technical writing grading training for the GTAs and UTAs in both FYEP tracks.

The purpose of this paper is to first describe the technical writing grading training that was implemented in our program, highlighting the changes to the training across three years. Second, we aim to provide an early evaluation as to the impact of this grading training program through both quantitative and qualitative assessment elements. In order to collect quantitative data, we examined two technical writing assignments – one lab memo and one lab report – that were unchanged across this three year period. These two assignments were analyzed to examine both the average score and standard deviation of the scores among all course sections for that assignment in order to assess the consistency of our grading. The qualitative element of our assessment consisted of focus groups with the TAs about the grading training process. These focus groups, conducted for both GTAs and UTAs, asked for feedback concerning what went well, what could be improved, what was most useful, and what additional topics or resources would be useful to implement going forward. We feel this feedback is critical in the development of a successful grading training program as we strive to ensure that the time we ask of our TAs for training is seen as beneficial, important, and worthwhile by both parties. Both the quantitative and qualitative assessment data are essential for understanding the impact of the training.

Finally, we present a “Tricks of the Trade” section that outlines some major themes that came out of our focus group discussions with TAs. This is provided in order to help others who are interested in implementing similar TA training programs at their institution.

Background
Our FYEP has used various techniques to train TAs in grading technical writing. In the following sections, we describe the process used in past years and the new approach implemented in Autumn 2014. Providing this information gives context for the results presented later in the paper. It should be noted that the training began in FEH but was later adapted to also be included in FE.

Previous Methods of TA Training for Grading Technical Writing
Prior to Autumn 2013, the grading training for FEH typically consisted of a single, two-hour session at the beginning of the semester where the TAs would work through the grading of a sample lab report in groups. The lead GTAs facilitated the training and answered questions. After this session, TAs were only approached again with a follow-up email or meeting if there were major anomalies noted in their grading.

Beginning in Autumn 2013, a new grading training effort, based on the TA training performed at Purdue9 to grade MEAs, was piloted in FEH. The first step in the process was to collect sample reports that represented a range of overall grades. A team of GTAs and faculty members graded the reports independently and had a meeting to discuss the scoring on the detailed grading rubric. Through this discussion, a baseline score and acceptable range was set for each grade category on each of the collected reports.
For the pilot grading training session, TAs were given instructions on grading for the semester and then worked with other TAs to grade a sample lab report. Afterwards, they were provided with a copy of the same sample report that was marked up with feedback to be used as an example. The TAs then received four writing assignments to grade as part of this training, which they graded independently and returned to the lead GTA. Their scores were compared to the baseline scores. Additionally, the feedback written on the report for the student was examined and judged based on quantity and quality. After collecting this information, each TA was given a copy of their spreadsheet with a feedback note about potential improvements they could focus on for grading in the future. The TAs graded the four writing assignments in sets of two, receiving feedback in between with the instructions to use the feedback to complete the grading of the final two assignments. Details about this feedback process, which remained unchanged, are provided in the “Implementation of New Grading Training Program” section.

Moving Towards New Training Methods for Grading Technical Writing

While this effort was a respectable first step toward a more comprehensive technical writing grading training, one missing component was an additional follow-up intervention for those that were still identified as grading either “too harshly” or “too leniently” after the last set of grading assignments was returned. Beyond the feedback provided – an email with their scores compared to a baseline and a short note on their grading – no additional measures were taken to help TAs further calibrate their scores and learn from the feedback. This missing piece could be a reason that the initial pilot did not result in a lower standard deviation in average report and memo grades between sections. Additionally, this version of the training was only implemented in FEH.

An additional enhancement to the pilot was to expand this training into the second course in the sequence in some way. While the initial pilot training was intended to impact both semesters, many employment and grading responsibilities change between semesters, thus it was possible that a number of TAs could enter the spring semester without any training in grading technical writing. Despite the second course having fewer lab writing assignments, there is a significant additional technical writing component added in the grading of the design project report. Therefore, it made sense to design the training program to cross both terms.

While this pilot was implemented in FEH, the TA training for the FE track consisted of a short lecture during orientation about technical communication and encouragement to attend a writing workshop conducted by the Engineering Communications within the College of Engineering. However, these workshops have had limited success due to attendance and TA turnover. Therefore this training improvement project sought to address this shortcoming and to establish a more consistent training procedure for TAs across the FYEP.

Implementation of New Grading Training Program

The updated and expanded pilot training program began in Autumn 2014. This project consisted of three phases. Each of these phases is described below to provide a detailed view of the enhanced TA training program for grading of technical writing.
Phase 1 – Preparation
Phase 1 of the project consisted of the initial preparation and planning for the grading training. This included finding representative technical writing samples, setting the baseline grades with a group of faculty and senior TAs, and creating sample “marked up” graded examples.

Phase 2 – Calibration Sessions and Feedback
Phase 2 of the project consisted of the initial training and calibration. This grading training occurred as part of the annual required TA training sessions. In addition to requiring the attendance of the GTAs and UTAs responsible for grading writing assignments, the faculty of the program were also encouraged to attend and participate. This was aimed at providing consistent exposure to the process and training throughout the FYEP. An overview of the grading process was given followed by an activity where the TAs graded in groups. After this sample grading activity all the TAs were provided a sample graded writing assignment to review and compare to.

After the in-person training session, TAs were given two assignments to grade. The TAs graded and returned the assignments and the scores were entered into a spreadsheet which compared their given scores to those of the established baseline scores. An example of this spreadsheet is shown in Table 1. Based on the average and standard deviation allowed from the baseline, the cells were formatted to be colored green if the grading is “too lenient” for a certain category and red if it is “too harsh”.

Table 1: Sample Calibration Spreadsheet

<table>
<thead>
<tr>
<th>Report 1</th>
<th>Average</th>
<th>Std. Dev.</th>
<th>GTA1</th>
<th>GTA2</th>
<th>GTA3</th>
<th>GTA4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Score</td>
<td>100</td>
<td>66</td>
<td>10</td>
<td>65</td>
<td>86</td>
<td>51</td>
</tr>
<tr>
<td>Content</td>
<td>48</td>
<td>5</td>
<td>44</td>
<td>61</td>
<td>38</td>
<td>52</td>
</tr>
<tr>
<td>Introduction</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Experimental Methodology</td>
<td>7</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Results and Description</td>
<td>10</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Discussion</td>
<td>10</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Summary and Conclusion</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Lab Specific Requirements</td>
<td>30</td>
<td>23</td>
<td>1</td>
<td>19</td>
<td>25</td>
<td>19</td>
</tr>
<tr>
<td>Language</td>
<td>9</td>
<td>2</td>
<td>10</td>
<td>12</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Complete, Concise, Clarity, Flow</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Professionalism, Tense, Person</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Spelling, Grammar, Punctuation</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Format</td>
<td>7</td>
<td>2</td>
<td>9</td>
<td>11</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Content Placement</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Formatting</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Labels, Referencing, Citation</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Survey</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Completed on Carmen</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
In addition to providing the scoring comparison spreadsheet shown in Table 1, TAs were also given notes concerning the amount and quality of feedback given for each assignment. For example, this feedback could have included:

“Your scores are in the appropriate range for each section of this report except the results section. You need to grade more harshly in the results section, as this report was missing key results and figures. Also, you included good comments throughout the report, but you need to provide more of these comments in the future.”

After receiving the personalized feedback, each TA graded two more assignments taking into account the feedback they received. Again, the scores were entered into a spreadsheet and personalized feedback was generated.

**Phase 3 – Follow-up and Monitoring**

The follow-up session had a different format in FE compared to FEH due to scheduling concerns. For FE, all TAs attended the additional session at a normally scheduled training time. For FEH only TAs that were identified as grading “too harshly” or “too easily” following the four grading assignments were required to attend an additional training session conducted by the lead GTAs and faculty. The session walked through one of the previously graded training assignments and detailed the baseline scores and the rationale behind them. The FEH TAs were then required to grade one additional assignment, and the same process as described in Phase 2 was used to provide feedback. A flowchart of Phase 2 and Phase 3 is shown in Figure 1.

![Figure 1: Flowchart of TA Training and Calibration Process](image)

In each subsequent year phases two and three will be repeated; however, that is beyond the scope of this paper. A flowchart of the process is shown in Figure 2.
Methods
Evaluating the training consisted of both a quantitative and a qualitative component. Approaching the evaluation in this way allowed us to gather both summative and formative feedback about the grading training. This knowledge will inform future improvements to the training program to ensure it continues to develop and remains effective for our needs.

The quantitative portion included grade tracking on writing assignments across sections of FE and FEH. Previous years grades were also included in this analysis, both before any grading training implementation and during the pilot year. While the grades for all lab reports, memos, and abstracts were recorded, only the data from FEH could be used to examine the effects of the initial pilot implementation and the subsequent implementation because the pilot did not occur in FE. Further complicating this quantitative comparison is the fact that the lab assignments and rubrics vary year to year. Sometimes this change can be the format of the assignment (report vs. memo vs. abstract) or it can be that it changes from a group to an individual assignment or vice versa. However, there were two writing assignments that did not change in the three years that were recorded – one lab report and one lab memo. These were the two assignments selected for the quantitative analysis. The analysis consisted of examining the variation and standard deviation of the scores among all course sections for that assignment. We recognize the limitation to this approach in that different students year to year completed these writing tasks. It is possible that the students could have had an effect on the results. With that in mind, we also implemented a qualitative component to our work.

The qualitative portion of our assessment included focus groups with TAs at the end of the fall term concerning the grading training process that they had completed. We followed the recommendations of Krueger and Casey to guide this portion of the evaluation. A total of three focus groups were held – one for FE GTAs, one for FE UTAs, and one for FEH GTAs (an FEH UTA focus group could not be held due to scheduling conflicts). During these focus groups, TAs were asked to provide information about what went well with the training, what could be improved, what was useful, and to comment on any additional possible topics or resources that could be implemented in the future to improve the grading training program. The focus groups were audio recorded and detailed notes were taken during each session that served as the data to be analyzed. We analyzed the qualitative data by allowing themes to emerge across the focus groups.
**Results and Discussion**

As described above, we collected both quantitative and qualitative data to inform the assessment of this training. Below we present and discuss the quantitative findings. The qualitative results were used to develop our “Tricks of the Trade.”

An initial quantitative comparison examined two assignments that were unchanged in the three years of data that was obtained. These two assignments also served as sample documents in the grading training and were part of the calibration phases. Table 2 shows the comparison of the FEH program average among the course sections and the standard deviation of the course section averages. While it appears that the initial Year 1 pilot did not result in lower standard deviations, the change from Year 1 to the current implementation in Year 2 did show improvement in both documents.

<table>
<thead>
<tr>
<th></th>
<th>Lab Memo</th>
<th></th>
<th>Lab Report</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Standard Deviation</td>
<td>Average</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td><strong>Year 0</strong></td>
<td>82.58</td>
<td>4.35</td>
<td>83.64</td>
<td>4.15</td>
</tr>
<tr>
<td><strong>Year 1</strong></td>
<td>84.29</td>
<td>6.03</td>
<td>80.30</td>
<td>5.00</td>
</tr>
<tr>
<td><strong>Year 2</strong></td>
<td>81.3</td>
<td>3.66</td>
<td>77.83</td>
<td>4.06</td>
</tr>
</tbody>
</table>

From this comparison of technical writing assignments, there was a clear improvement in the standard deviation for grades across the program between Year 1 and Year 2. For lab memos, the standard deviation for grading dropped from 6.03 to 3.66, and for lab reports from 5.00 to 4.06. The standard deviation values for Year 2 were also lower than those seen in Year 0. Additionally, the average scores for both memos and reports were the lowest in Year 2. Figure 3 (a) and (b) shows the three years of course section averages for the lab memo and lab report, respectively. From here the improvement in the lab memo distribution from Year 1 to Year 2 is clearly shown. The program average and +/- 1 standard deviation is displayed. Each data point corresponds to a course section average for that assignment. The data depicted in Figure 3 clearly demonstrates an improvement (i.e. decrease) in the standard deviation between Year 1 and Year 2 for both technical writing assignments. The average grades for these assignments also decreased slightly across the program.
This analysis of the technical writing assignment grades highlights an improvement in the standard deviation for grading across our FYEP between Year 1 and Year 2 following the implementation of our expanded grading training program. One possible explanation for this improvement is that the TAs responsible for grading these technical writing assignments have been provided with a substantial increase in the amount of practice in grading these writing assignments. The addition of the four grading assignments provided as part of their training afforded the TAs with significant opportunities to practice their grading of technical writing, as well as offering them feedback as they progressed.

An additional trend revealed from the data presented is a decrease in the average scores across all sections. The averages for Year 2 are lower than both Year 1 and Year 0 for all cases. We do not believe that this trend represents a difference in the technical writing aptitude of our first-year engineering students but rather indicates that our expanded training program has better prepared our TAs to correctly identify and respond to students’ mistakes. This data may indicate a better understanding of both the rubrics and grading policies across the program that will continue to normalize over time.

“Tricks of the Trade”
Based on the focus groups conducted with the UTAs and GTAs, the following themes emerged from the discussions. We have framed these themes as “Tricks of the Trade” to help others who...
are interested in implementing a similar program at their institution. Based on our focus groups, the TAs:

- **Appreciated receiving feedback on grading** – All the groups mentioned that they liked receiving feedback on the grading training. Some groups mentioned that they would prefer that this feedback was more descriptive. *We recommend that all programs provide feedback to TAs both on and throughout the training process.*

- **Liked being compared to a standard** – The GTA groups mentioned that they liked being compared to a standard. Also, since these GTAs are overseeing UTAs, they also mentioned that they liked knowing that the UTAs had seen the standard as well. *We recommend that expectations are communicated with all members of the teaching team so that everyone approaches grading in a similar fashion.*

- **Acknowledged differences due to experience of TAs** – The groups mentioned that the varying levels of experience in the TAs impacted the grading training. There were suggestions made to change the training based on previous experience. It was also mentioned that those TAs with previous experience knew the grading expectations ahead of time and it impacted the training process. *We recommend developing different types of training based on experience or involving those with experience as mentors in the training process.*

- **Suggested changes to the training schedule** – The TAs all indicated that the grading training could be moved earlier in the semester and possibly completed before the annual orientation. Because the TAs also indicated that the training took a lot of time, this may mitigate the impact of the time and effort required. *We recommend critically evaluating when the training occurs based on your university’s academic schedule and making adjustment as needed to ensure that the TAs are not overburdened by training tasks.*

- **Recognized the seriousness of the training** – The FE TAs mentioned that this grading training had a level of seriousness that was not seen in other trainings offered and required in the program. They mentioned being surprised that the training assignments that were turned in were looked at because this had not happened in the past. This seriousness is related to the feedback on the grading training. If the training is required, it is important for everyone involved to take it seriously so that TAs complete the training to the best of their ability and are able to get the most out of their effort. *We recommend that everyone involved in the training sees it as a priority, and in turn helps transfer that sense of priority to the TAs who will be completing the training and ultimately the grading.*

The “Tricks of the Trade,” which are in italics above, highlight the salient themes from the focus groups. We intend for these statements to serve as recommendations to others and plan to conduct further forms of assessment to be able improve our program in the future. These are not the only things that can be done to have a successful grading training program, but they are a starting place for those interested in enhancing the training of their TAs and ultimately improving the educational experiences of their students.
Conclusion
This work presents the implementation of an expanded grading training program for TAs in a FYEP with the goal of improving the grading of technical writing assignments. Technical communication is recognized as a highly valued skill for engineers in all disciplines, and with the increasing amount of focus placed on this skill in the classroom, it was necessary that we improved the way in which we prepared our TAs to grade and provide feedback for these writing assignments. We have highlighted above the changes made in our training program over the past three years. This training has been expanded to include multiple phases at the beginning of each school year, illustrated in Figures 1 and 2 above, with significant increases to both the amount of grading practice provided as well as the amount of feedback offered.

The impact of this new grading training program was assessed using both quantitative and qualitative elements in order to obtain an early evaluation of these efforts. Quantitative analysis demonstrated a decrease in the standard deviation for grading across our FYEP between Year 1 and Year 2 following the implementation of our expanded grading training program, which we think may indicate a better understanding of rubrics and an improved ability of our TAs to identify and respond to student mistakes. Qualitative results were gathered by holding TA focus groups, and the feedback from these sessions led to the “Tricks of the Trade” section above. In short, our TAs appreciated the additional feedback provided as well as having a standard to compare their grading too. It was also clear that the expanded training demands a significant time investment, and thus it is important to evaluate the timeline of the training based on the university’s academic schedule to avoid overburdening the TAs.

Based on these results there are potential areas for future improvements and assessments of the grading training. Additional analysis of the training could include an examination of the feedback provided both on training documents and on the returned assignments to see if the increased training leads to better quality feedback provided to the students. Further improvements to the training process will be the implementation of an electronic method to collect the training scores and provide real-time feedback. Feedback on the level of comments written on the training document would still need to be done by hand, but automating the quantitative collection and feedback will minimize the resources required to manage the training process and expedite the important feedback loop to the TAs. We plan to move the training to begin before the semester to allow the TAs additional time to work on the training materials. We also plan to include additional tracking of grades throughout the semester to provide more training to TAs who may require it. Additionally, we also want to continue to collect qualitative feedback from the TAs throughout this process so we can continue to make improvements to make the training as useful as possible.

Acknowledgements
We would like to thank the TAs who participated in the grading training and focus groups. We would also like to thank the TAs who helped develop materials for the training, especially Beth Yoak and Russ Stech who helped facilitate this process. Additionally, we would like to thank Monica Okon for her help in conducting the focus groups.
References


