Correlation Between Individual Grading Metrics and Final Course Grades in Senior Capstone Design: Indicators of Success

Dr. Tristan Martin Ericson, York College of Pennsylvania

Dr. Tristan M. Ericson is an Associate Professor at York College of Pennsylvania and has been teaching there since 2013. Prior to this appointment he was a Visiting Professor at Bucknell University. His teaching interests include solid mechanics, vibrations, materials science, and MATLAB. He advises the YC Racing Formula SAE team. His technical research interests include vibrations of planetary gear systems, strengthening 3D printed materials, and making things go faster. He enjoys activities that promote STEM fields in local high schools. Hobbies include 70s Japanese motorcycles, vintage machinery, and rural living. He received his PhD from Ohio State University in 2012.

Correlation Between Individual Grading Metrics and Final Course Grades in Senior Capstone Design: Indicators of Success

Abstract

Students need feedback about their performance throughout any course, but particularly in capstone design because (1) a large team project is different from lectures and labs, (2) a failure in a two-course sequence is more costly, and (3) the skillset for success is different from other courses. A significant portion of a typical capstone course grade is normally evaluated at or near the end of the term, whether in the form of a final report, completed prototype, or summative design presentation. This paper evaluates how well certain grading metrics – for which students *do* receive feedback throughout the course – predict a student's final grade.

The metrics used to evaluate final course grades in a Capstone Design project at York College of Pennsylvania are: (1) brief periodic stand-up presentations, (2) "milestone" tasks completed according to an instructor-defined timeline, (3) written summative technical report, (4) poster created within project sub-teams, and (5) professionalism evaluated by both instructor and peers. Students receive feedback throughout the courses from the first and second metrics listed. The third and fourth are end-of-semester deliverables. The last is typically evaluated at the end of the semester as well. Data is available and analyzed across eight years with a consistent project (Formula SAE), grading system, and instructor. There are typically about 20 students, distributed across sub-teams of three to five members, leading to a total of 164 individuals in the data set.

Results show that the first two metrics above, which provide students feedback throughout the course, are reasonably good tools for predicting final grades: they are ranked second and third out of the five measures used. In particular, periodic standup presentation grades (metric 1) predict the final course grade within one half letter grade 84% of the time and "milestone" tasks (metric 2) do so 85% of the time. This is better than the correlation of two end-of semester deliverables: metric 3 at 70% and metric 4 at 69%. The best indicator, however, of final grade status is professionalism (metric 5) at 93%. Professionalism has typically not been evaluated until the end of the semester. Fortunately, it is easy to conduct some professor and peer evaluation earlier in the term to provide students that valuable feedback. In addition to the method just described – prediction within one half letter grade – this paper uses several techniques to measure the correlation of grading metrics against final course grades and evaluate whether students are receiving adequate feedback throughout the semester.

1. Introduction

Expectations and essential skillsets needed to succeed in capstone design may differ from other courses in engineering curricula. At York College of Pennsylvania, capstone is less-structured, providing students autonomy to budget their own time. Occasionally, analytically strong students struggle and weaker ones thrive in the intensive project-based setting. Students are evaluated differently too. York College capstone design students are not assigned weekly homework, nor do they take tests/quizzes based on course material (lectures, readings, etc.) as in prior courses. Instead, throughout the semester we evaluate their notebook records, stand-up presentations, and individualized "milestone" deliverables according to a schedule.

The fundamental question I'm asking is: *Are my students getting sufficient feedback throughout the course to know where they stand in the class?* I'm asking this question because students receive consistent feedback on two grading metrics throughout the course but other grading metrics are naturally evaluated at the end of the course. Therefore, the question at hand becomes: *Are the grading metrics that provide feedback to students throughout the course reasonably predictive of the final grades?* Whether the answer is *yes* or *no*, the key follow-up question is: *Can anything be done to help students better understand their status in the course?* These grading metrics (with my terminology in parentheses), are:

- Brief oral presentations to the team with notebook review (tri-weekly presentations)
- Key tasks completed according to a prescribed timeline (milestones)
- Summary end-of-semester technical report (tech report)
- Posters created within sub-teams (posters)
- Overall contribution to the team success and as a good citizen (professionalism)

The first two metrics listed provide feedback to students throughout the semester, but the other three are typically evaluated at the end of the semester.

Capstone design is a unique course, and the grading metrics often differ from others in the curriculum. Project work is common, so instructors may need to find a balance between group and individual assignments [1] because while the efficiency of assigning group grades is attractive, that strategy is not granular enough to recognize individual contributions [2]. Some professors are not fully confident in their ability to create assignments and assess them in a capstone course [3]. Once appropriate assignments are developed it may be difficult to ensure consistent grading across projects and fair assignment of credit, but several authors have developed systems that provide equity [4]. Peer evaluations are commonly used to this end [5] and some instructors utilize online systems [6]. Midterm peer ranking may motivate underperforming students [7]. Peer evaluations generally correlate well with traditional metrics, like exams [8]. Student feedback indicates that complex rubrics are not necessary to evaluate peers [9]. Even from an instructor perspective, the unfamiliar labyrinth of capstone design assignments and assessment is daunting. I'm wondering amidst my own personal experience: Are my students getting enough information through the course to know where they stand?

2. Senior Capstone Design at York College of Pennsylvania

A brief overview of the Capstone Design sequence in the Mechanical Engineering program at York College is appropriate, however, readers will likely notice marginal differences between what is presented here and any familiar program. Senior students in our engineering programs take a two-semester design course sequence in their last two terms in the program: Capstone Design I and Capstone Design II. The first course focuses on research, defining specifications and the scope of work, iterative design, analysis, and a final detailed design proposal containing solid models, drawings, FEA, and any key piece of evidence necessary to defend the final design. In the second course students must build, test, and evaluate their product with appropriate testing (along with any necessary design modifications identified through this phase).

Naturally there are several prerequisite courses for Capstone Design I: System Modeling and Control, Instrumentation and Microprocessor Lab, Fluid Mechanics, and Machine Design. Other important sophomore/junior level courses are often prerequisites for one of these. We regard Machine Design as the most important, and one essential prerequisite on the list. In that course, and along with theoretical content, students complete a design and build project on teams of three to four members. Because of this and our required co-op program where students typically work closely with a couple engineers, we generally desire larger senior capstone teams so that students gain experience in that environment. Our capstone teams usually contain six to ten students. I advise the largest project, our Formula SAE team, which typically has 20-25 mechanical, electrical, and sometimes computer engineering students.

3. Scope of Inquiry

The Formula SAE project tends to have a higher rate of failure than others, and we attribute this observation, largely, to the dynamics associated with a higher number of students on the team. We want to reduce the likelihood of course failure, so this work evaluates whether the feedback students receive throughout the semester reasonably predicts final course grades. Failures occur more commonly in the first course of the sequence, so this work will analyze the data associated with Capstone Design I. The analysis is limited to the Formula SAE project because it contains the subset of capstone students with a higher failure rate compared to other projects. Additionally, grade data exists for this project under the same advisor, applying the same grading rubrics for 164 students across eight years of the project. That is the data set for this work.

4. Course Grading Metrics

Five metrics are used to evaluate final grades in both capstone design courses in the sequence. These metrics reflect tasks expected of professional engineers on a real-world engineering design team. Each one comprises strong components of communication and technical contribution. Table 1 identifies the five metrics and their weights in calculating final course grades.

Table 1: Capstone Design I grading metrics and their percentage of overall course grade.

Metric	Weight
Tri-Weekly Presentations	**20%**
Milestones	**30%**
Final Technical Report	**20%**
Sub-team Poster	**5%**
Professionalism	**25%**

The two principal metrics yielding instructor feedback throughout the semester are the *Tri-Weekly Presentations* and *Milestones*. We desire one or both of these to be strong indicators of success in the course and highly correlated to the overall final course grade. By nature of the assignments, the *Final Technical Report* (Tech Report) and *Sub-team Poster* are submitted at the end of the course, so there is no meaningful opportunity for students to react to feedback on those assignments. *Professionalism* is typically evaluated at the end of the term, but there would be opportunities to assess it earlier as well, such as at the semester midpoint.

4.1. Tri-Weekly Presentations (throughout semester)

Engineers often present progress reports to superiors or a team of peers. The ability to document one's own work and organize it in a way that the group easily understands are crucial skills. Additionally, professionals should be able to respond to technical questions with coherent, succinct, and sufficient answers. The tri-weekly presentations assess these skills.
Once every three weeks an individual student will discuss their progress since the prior report using PowerPoint, making sure to address: planned tasks since the last cycle, progress during the current period, and a task list for the coming three-week period. The presentation should be about 5-6 minutes with an additional 2-3 minutes for questions. Students are required to keep a hardbound engineering notebook throughout the duration of the project, and the instructor reviews and writes comments in the notebook during the presentation. I typically expect 20-30 pages of notebook content per reporting period. Each student keeps a grade log in the notebook used by the instructor to report back grades for the three-week progress period. The notebook is returned with these grades and any comments after the presentation, so students receive immediate quantitative and qualitative feedback. The whole of the tri-weekly presentation, including notebook content is graded on a 10-point scale according to Table 2. Each student delivers three presentations in Capstone Design I and four in Capstone Design II. Presentation scores are averaged across the semester to yield the score for this first grading metric.

Table 2: Grade indicators and point valuation for each tri-weekly presentation.

	Points
Notebook content	**3**
Presentation quality	**2**
Overall progress	**5**
Total	**10**

4.2. Milestones (throughout semester)

The success of any large-scale team project (such as Formula SAE) depends on a breakdown of the overall objective (designing and building a capable racecar) into smaller tasks completed according to a timeline. The Formula SAE team at York College of Pennsylvania is associated with the capstone design sequence, so it's comprised entirely of first-year members. A team of 20-25 new students isn't likely to have the experience and foresight to successfully plan a task list and timeline for the entire project, so this outline is largely provided by the instructor. The entire student team is divided into six sub-teams (frame, suspension, engine and drivetrain, body and aero, electrical, and systems and data acquisition) and each sub-team will have about 10-15 tasks, or "milestones", identified with delivery dates. This level of detail is sufficient to keep the project on schedule while allowing students some autonomy to identify and plan the sub-tasks necessary to meet the deadline for each milestone. Examples of a few milestones (the points system is explained below) are:

- Control arms designed/analyzed and modeled in SolidWorks (4 pts.)
- Preliminary front/rear chassis members designed/analyzed (3 pts.)
- System designed to display gear position (2 pts.)
- Engine oil system components designed/chosen/analyzed (4 pts.)

The quality of student work is evaluated for each milestone according to a rubric with four indicators: research quality, analysis, degree-of-completion, and confidence. Milestones are submitted as two-per-page printouts of PowerPoint slides because the I generally desire pictures, graphs, and bullet lists. PowerPoint is well-suited for this format, but students do not give a stand-up presentation to the class. They are, however, typically asked to brief the team on the outcome of their milestone assignment in a round-table discussion.

Aside from evaluating the quality of milestone submissions, it is necessary to assess the overall volume of milestone work that each student contributes to the project. It would, for example, be insufficient to base the total course milestone grade one very well executed but relatively small task. To this end, each milestone has some number of points (in the bullet list above) that is associated with the relative workload expected for that task. The points are allocated with the expectation that an average student will earn 10 points throughout the semester. This means that if a single student works on the "control arms designed/analyzed and modeled in SolidWorks" milestone above, that would represent 40% of an expected semester workload. Multiple students are welcome to collaborate on a milestone and split the points proportionally. (The instructor has the final say on point allocation based on what each student contributed.) The final milestone course grade is calculated by a linear combination of grades and points for each milestone, normalized to the 10 expected points. In this way a student earning more/less than 10 points will see a final grade increase/decrease appropriately. A more detailed explanation of the milestone system with a grading rubric and sample grade calculations is given in [10].

4.3. Final Technical Report (end of semester)

There is likely nothing novel about the final tech report deliverable or its evaluation compared to capstone design courses at other institutions. It is the final summative report that each student writes to present their entire semester's work. These reports are used to evaluate a student's overall technical contribution and ability to communicate that work effectively by written report. They also a critical resource for future teams seeking to understand what's expected in capstone design and acquire a baseline understanding of prior work.

Tech reports are evaluated using a rubric based on several ABET outcomes: (2) ability to apply engineering design, (3) ability to communicate, (4) ability to recognize ethical and professional responsibilities, (7) ability to acquire and apply new knowledge, and other attributes of general engineering design. Each of these areas just listed has three to five indicators that are evaluated in the report rubric. Students are provided the rubric indicators prior to the tech report due date.

4.4. Sub-team Poster (end of semester)

Students in any capstone project at York College will normally present their work at the end of each semester to a variety of stakeholders: industry professionals, alumni, project clients, etc. This is typically accomplished through an expo poster session demonstrating the design and analysis (first semester) and final product (second semester) to the audience. Formula students create posters within their sub-teams, and typically one or two posters are necessary within each of the six groups. Of course, this team will also present their work to judges at the May FSAE competition in Michigan, and posters have often been utilized for this purpose too.

Posters are graded on a simple rubric evaluating (1) technical content and (2) visual presentation. Grades are assigned according to poster, so they typically map to 2-4 different students. Table 1 shows that this metric carries the least weight on the final grade, and I expect that the poster score will generally help (at least not harm) the students' final grades.

4.5. Professionalism (typically end of semester)

So much of what makes an engineer a valued contributor in the workplace is difficult to quantify in most grading metrics, even those above that are designed for such a course as this. Valuable soft skills are often demonstrated (or not) out of the instructor's (or supervisor's) field of vision. Even those that are evident to the instructor are not necessarily captured by the metrics above. Yet these skills are critical to keeping a project on task and fostering a positive culture within the team. The professionalism metric seeks to acknowledge the efforts of students who:

- Demonstrate reliability and dependability with their work
- Approach the work of the team with a positive attitude
- Prioritize success of the team by helping others
- Rework a milestone if needed, even after it was submitted and graded

The professionalism grade is evaluated by combining an instructor observation score with peer evaluations. Those two indicators are typically weighted equally and evaluated at the end of the semester. Occasionally, and typically in the second semester, a midpoint evaluation (instructor and/or peer) is used to provide formative feedback to the students, and those results may be included in the final evaluation. Instructor observation is a simple percentage.

Numeric scores are not collected from students in the peer observation. As in [5] and [8], students are asked to identify the quality of their peers' work in one of nine categories. For example, a teammate's work is "Very Good" if they "consistently did what he/she was supposed to do, very well prepared and cooperative". Once feedback is collected, each description is converted to a score ("Very Good" relates to 87.5%) and averaged with the feedback score from each team member. Every member of the entire team is invited to evaluate every other member, but they are free to leave any member unrated if they don't know enough to evaluate that student. In practice, about 20% of individual evaluations are left blank, that is, I receive data from about 80% of the team for an average student. I have consistently found that the peer evaluation scores are remarkably close to the instructor observation scores and similar to scores generated by the other metrics. Table 3 shows the average score across all 164 students in the eight-year dataset. It may seem surprising that the numbers are so similar. In fact, the more detailed analysis below will show that the professionalism metric is the best indicator of overall course grades.

Table 3: Average grades of all 164 FSAE capstone students throughout eight years.

	Score
Instructor Observation	**84.0%**
Peer Evaluation	**83.4%**
Combined Professionalism	**83.7%**
Average Course	**83.6%**

5. Evaluation of Grading Metrics

My goal is to determine which grading metrics above are most predictive of overall course grades and see if the two metrics that provide feedback throughout the semester – *Tri-Weekly Presentations*, and *Milestones* – are sufficient for this purpose. This evaluation uses four methods described below. These techniques are not significantly different from one another, but they provide different perspectives of the data and point to common conclusions.

5.1. Method 1: Overall Difference Between Metrics and Final Course Grades

The first method looks at the difference between every student's grade for each metric compared to their final course grade. A low (near zero) difference indicates that the metric's grade correlates well with the final course grade. Additionally, a low standard deviation of the difference for each metric is desirable. A low difference and low standard deviation of differences would indicate a good predictor. The box plot in Figure 1 shows the overall difference, averaged across all students in the dataset, for each metric compared to the final course grade, with the standard deviation.

The data implies that the professionalism score is most predictive of the final course grade, having the lowest average difference and lowest standard deviation in the data. Although the tech report has the second-lowest standard deviation, it has the highest average difference. Therefore, the tri-weekly presentations and milestones – those from which the students receive feedback throughout the semester – appear to be the metrics with the second and third best correlation.

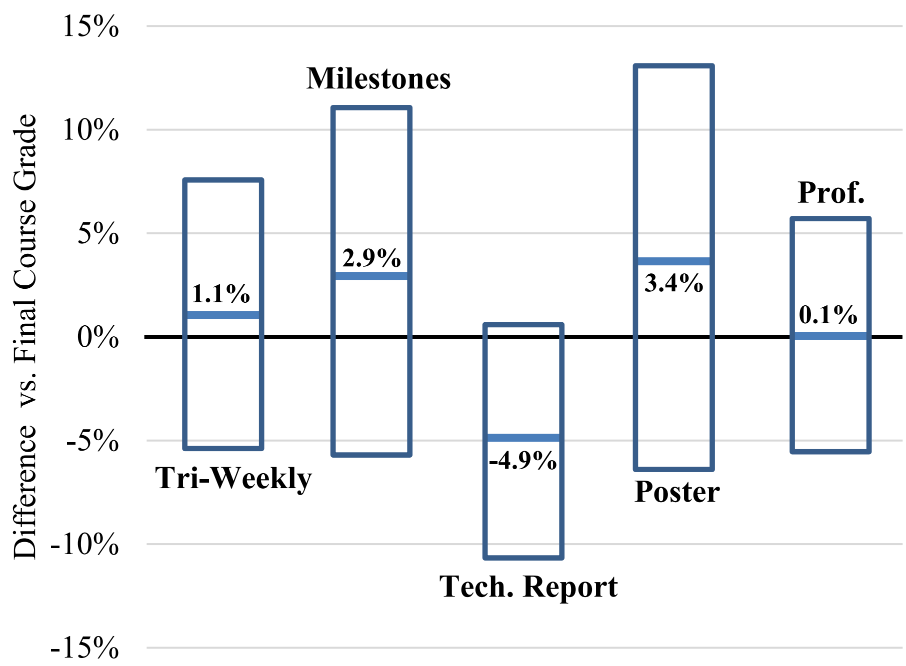

Figure 1: Box plot with average difference between each grading metric and final course grade along with standard deviation for all 164 studnets across eight years.

The data in Figure 1 is rolled up across quite a few years and many students. We are at risk of overgeneralizing the results and missing some meaningful variations by year or by student performance (grade level). Tables 4 and 5 address this. Table 4 expands the data from these figures out across each year in the dataset, and Table 5 expands the data out by assigned course "letter" grade. Grades 2-4 are *passing* and a 1 is *failing*. Students could earn a 0 in the course, which is also *failing* but that data is removed because these students typically don't submit the final tech report, and that skews the data. Those few students clearly understand that they are failing the course. Our focus is on the students earning a 1 and whether they understand that they are in danger of failing. Low numbers (white) in Tables 4 and 5 are good and high ones (red) are undesirable. (This color scheme is used throughout the paper: whiter backgrounds are good and redder are bad.) It is apparent that – across all years and grade levels – the professionalism and tri-weekly presentation metrics best correlated, respectively, to the final course grade.

Table 4: Average difference between each metric and the final course grade (left) and standard deviation of those differences (right) expanded out by course year.

Average Difference						Standard Deviation				
Tri-Weekly	**Milestn**	**Tech. Report**	**Poster**	**Prof.**		**Tri-Weekly**	**Milestn**	**Tech. Report**	**Poster**	**Prof.**
1.1%	**2.9%**	**-4.9%**	**3.4%**	**0.1%**	**Total**	**6.4%**	**8.7%**	**5.7%**	**9.6%**	**4.8%**
0.7%	2.9%	-4.5%	0.2%	-0.6%	**2024**	5.6%	6.5%	5.8%	6.7%	4.3%
2.4%	5.0%	-5.7%	2.4%	-1.0%	**2023**	6.3%	7.9%	5.5%	10.1%	5.1%
2.2%	3.7%	-6.9%	2.2%	-0.3%	**2022**	5.4%	9.6%	5.5%	8.4%	4.8%
0.9%	0.9%	-5.7%	10.2%	2.4%	**2021**	8.9%	11.1%	5.5%	11.2%	4.9%
1.2%	4.7%	-4.3%	1.8%	-2.2%	**2019**	5.6%	8.4%	5.4%	9.7%	4.8%
-1.2%	3.2%	-2.0%	2.0%	-0.9%	**2018**	4.9%	8.5%	6.1%	9.2%	3.8%
-0.2%	-0.7%	-2.6%	5.9%	3.6%	**2017**	7.6%	10.1%	6.4%	9.4%	4.7%
2.4%	2.6%	-6.8%	1.1%	0.6%	**2016**	5.8%	4.5%	4.9%	7.2%	3.2%

Table 5: Average difference between each metric and the final course grade (left) and standard deviation of those differences (right) expanded out by final letter grade.

Average Difference						Standard Deviation				
Tri-Weekly	**Milestn**	**Tech. Report**	**Poster**	**Prof.**		**Tri-Weekly**	**Milestn**	**Tech. Report**	**Poster**	**Prof.**
1.1%	**2.9%**	**-4.9%**	**3.4%**	**0.1%**	**Total**	**6.4%**	**8.7%**	**5.8%**	**9.6%**	**4.8%**
0.0%	4.8%	-6.2%	-3.1%	0.6%	**4**	4.3%	6.2%	4.8%	5.5%	3.1%
0.2%	5.2%	-6.3%	0.6%	-0.5%	**3.5**	6.2%	7.8%	4.7%	5.5%	3.3%
0.4%	4.9%	-6.0%	3.7%	-1.0%	**3**	5.5%	7.5%	6.1%	4.8%	5.0%
-0.9%	-0.3%	-2.0%	7.1%	1.1%	**2.5**	6.2%	9.5%	5.9%	4.1%	6.3%
4.5%	-4.6%	-1.8%	14.5%	-0.4%	**2**	8.1%	11.3%	6.2%	6.4%	6.6%
2.7%	-6.3%	2.2%	18.7%	-1.6%	**1**	5.9%	10.0%	6.5%	6.6%	8.5%

5.2. Method 2: Prediction of Final Course Letter Grade

Beyond looking at the average difference between the final course grade and each metric (method 1), we are essentially interested in whether the assigned "letter" grade that a student receives is reflected by the letter grade that would be predicted by looking at any one metric alone. Table 6 shows how the calculated final course percentage (according to Table 1) aligns with the assigned course grade. A final course grade of 1 or 0 is failing.

Table 6: Final course grade scale.

Calculated	Assigned
>90%	4
85% - 90%	3.5
80% - 85%	3
75% - 80%	2.5
70% - 75%	2
60% - 70%	1
< 60%	0

Figure 2 shows the data graphically with scatter plots for each student's metric grade plotted against the final letter grade they were assigned. The green boxes represent areas which the final letter grade would be predicted exactly by the grade percentage of each metric. The gray dashed boxes define areas which the metric predicts the final course grade within a half of a letter grade. More data points in these areas are better.

The data for metrics predicting the final assigned grade exactly isn't great. When the overall grade, for example, is 89% and the metric is at 91%, a data point is outside a green box, but the difference in these values is quite small. If, however, a data point is inside the gray dashed boxes, the metric is predicting the final grade reasonably well, considering the prediction is within half a letter grade. As before, professionalism is the best predictor, and the tri-weekly and milestone metrics would follow (better than the tech report and poster metrics).

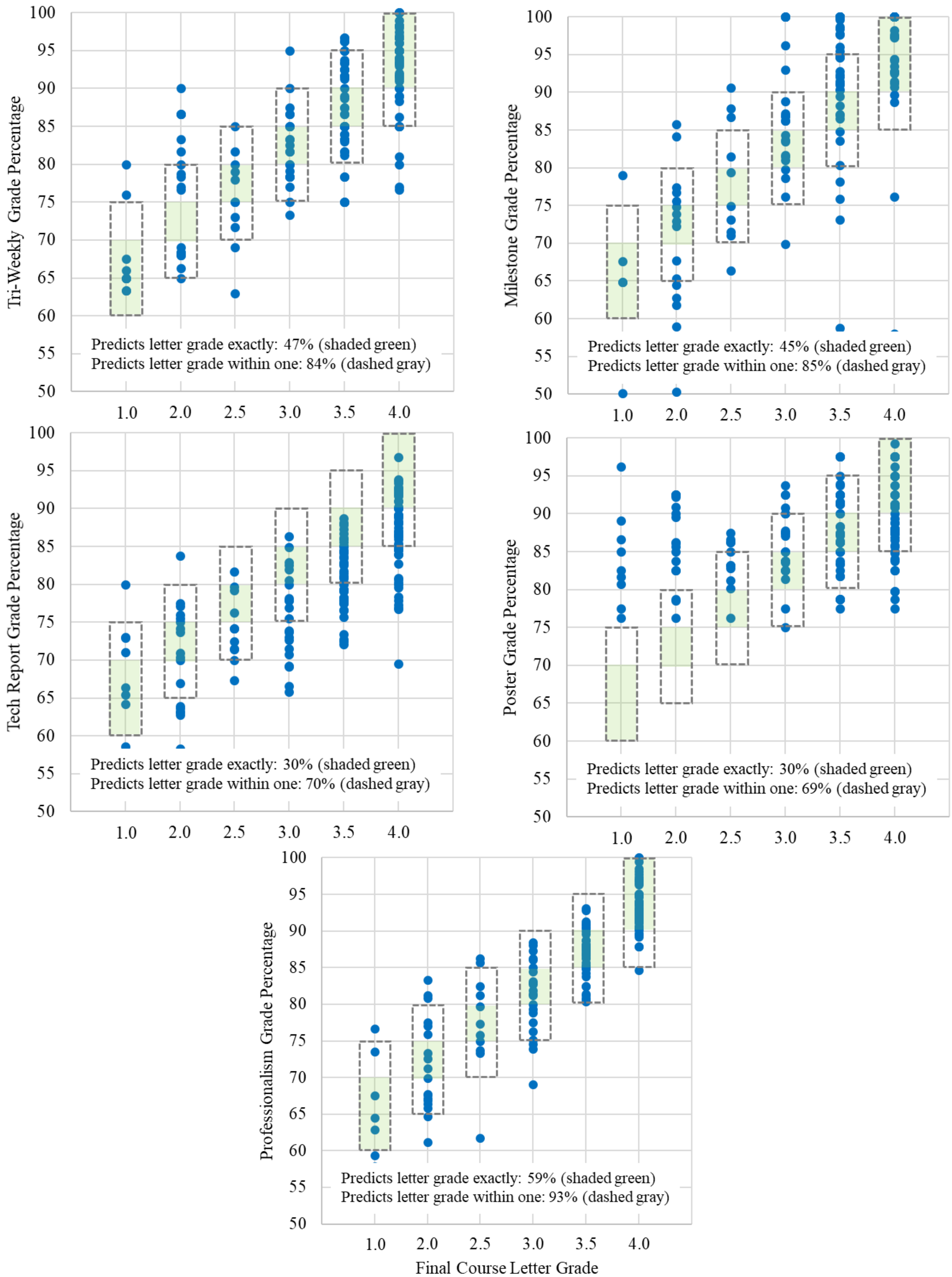

Figure 2: Scatter plots of all students' grade by metric vs. final course letter grade.

5.3. Method 3: Closeness to Final Course Percentage

A similar approach to that presented directly above would be to avoid the discrete 4-through-1 letter grade bins and look at percentage of occasions which a metric predicts the final course grade within 5% (half letter grade demarcation) or 10% (one letter grade). Table 7 does just this. We can say, again, that professionalism is the most predictive metric, and this time the tri-weekly presentations are clearly a second-best indicator.

Table 7: Percentage of instances which the final assigned letter grade would be predicted within 5% (left) and within 10% (right).

Prediction within 5%						Prediction within 10%				
Tri-Weekly	**Milestns**	**Tech. Report**	**Poster**	**Prof.**		**Tri-Weekly**	**Milestns**	**Tech. Report**	**Poster**	**Prof.**
64%	**38%**	**43%**	**49%**	**77%**	**Total**	**91%**	**71%**	**80%**	**78%**	**96%**
82%	43%	39%	59%	91%	**4**	96%	89%	79%	88%	98%
57%	27%	36%	66%	84%	**3.5**	89%	57%	75%	93%	100%
64%	55%	41%	55%	64%	**3**	95%	68%	73%	91%	95%
64%	36%	55%	36%	73%	**2.5**	91%	64%	100%	82%	91%
28%	28%	56%	6%	50%	**2**	72%	67%	94%	28%	94%
67%	33%	56%	0%	44%	**1**	89%	56%	78%	11%	78%

5.4. Method 4: Regression Analysis

Finally, we plot every student's grade within a metric vs. their overall course grade and consider the slope and R^2 value of a linear trendline. Figure 3 provides this information for each grading metric. We desire a slope of 1 and high R^2 value to indicate that a metric is highly predictive of the final course grade. Again, the professionalism score is the best metric with a slope of 1.01 and $R^2 = 0.79$. The next best metric is the tri-weekly presentation grade, followed by either the milestones or tech report, depending on whether slope or R^2 is more valued. The poster score (a grade always shared across a few teammates) is by far the worst predictor, showing poor correlation.

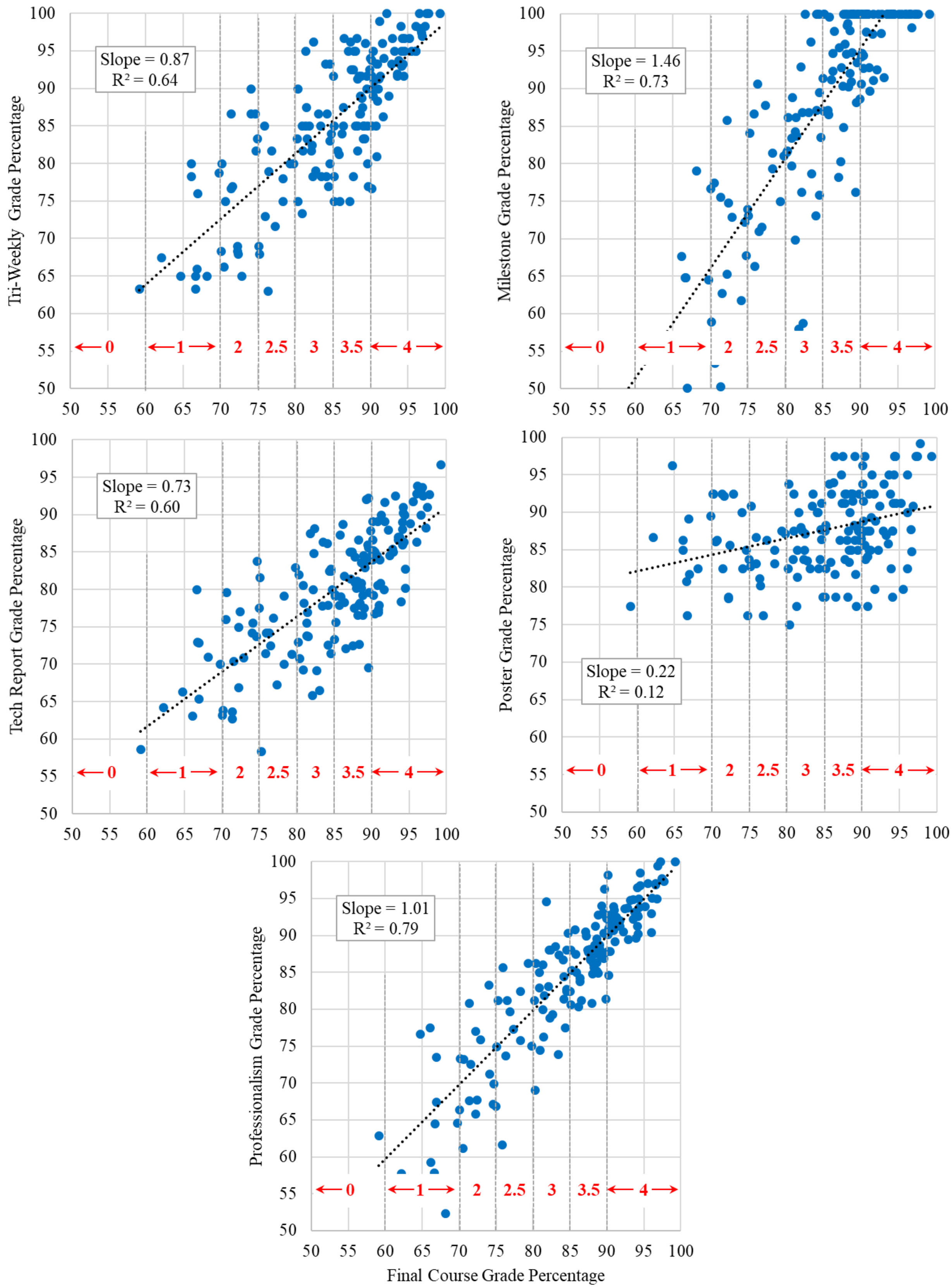

Figure 3: Scatter plots of all students' grade by metric vs. final course grade percentage.

6. Evaluation of Metrics

The professionalism metric is the best predictor of final course grade among all methods used. The tri-weekly presentations are also consistently the second-best metric, although it is essentially tied for second with milestones in method 2. In most metrics the milestone metric scores better than the tech report, though there is not much separation. The poster metric tends to have the weakest correlation. Table 8 summarizes the results from the four methods presented and ranks them according to their apparent effectiveness in predicting the final course grade.

Table 8: Summary of the results from the four methods of characterizing individual grading metric correlation with final course grade.

Method		Tri-Weekly	Milestones	Tech. Report	Poster	Prof.
1)	Mean Difference	1.1%	2.9%	-4.9%	3.4%	0.1%
	Standard Deviation	6.4%	8.7%	5.7%	9.6%	4.8%
2)	Matches Final Letter Grade	47%	45%	30%	30%	59%
	Within One Half Letter Grade	84%	85%	70%	69%	93%
3)	Within 5% of Final Course	64%	38%	43%	49%	77%
	Within 10% of Final Course	91%	71%	80%	78%	96%
4)	Regression Analysis Slope	0.87	1.46	0.73	0.22	1.01
	R^2	0.64	0.73	0.60	0.12	0.79
	Metric Rank:	#2	#3	#4	#5	#1

7. Breaking Down Professionalism

Section 4.5 noted that professionalism is a combination of peer evaluation and instructor observation. Anecdotally, these indicators have been fairly similar over time. But because professionalism is the most predictive metric, it would be valuable to see if either instructor observation or peer evaluations are a better predictor of the overall course grade. Table 9 shows the indicators of correlation for instructor observation, peer evaluations, and the combined professionalism score from methods 1 and 4 in this analysis. All three indicators are good, but the combined professionalism score is actually better than either instructor observation (scores tend a little higher) or peer observation (scores tend a little lower) alone.

Table 9: Indicators of correlation with final course grade from methods 1 and 4 for the combined professionalism metric and its two contributing components.

	Instructor	Peer	Combined
Mean Difference	**0.3%**	**-0.2%**	**0.1%**
Standard Deviation	**4.7%**	**6.4%**	**4.8%**
Slope	**0.97**	**1.07**	**1.01**
R^2	**0.79**	**0.69**	**0.79**

8. Reflections and Conclusions

My goal has been to determine if the tri-weekly presentations and milestones – the two grading metrics from which students receive consistent feedback through the semester – are predictive of the final overall course grade. These metrics are within 10% of the final course percentage 91% and 71% of the time, respectively (Table 8). Additionally, they predict the final letter grade within one half increment (e.g., a 2.5 vs. 3.0) in 84% and 85% of instances, respectively. These rates are fairly predictive of the final course grade, and notably better than the tech report or poster metrics, those that necessarily come at the end of the semester. They are not, however, better than the professionalism metric, which stands out as an excellent predictor of final course grade. Although professionalism is often evaluated at the end of the semester, there is opportunity to include a midterm assessment so students can get some of this feedback during the semester. This work provides strong evidence for doing so in the effort to inform students about their overall grade status in the course.

Experience over the years also suggests that students should be reminded to pay closer attention to their grades as they receive feedback. The LMS course associated with the Formula SAE project is non-academic because students from various disciplines (and different capstone design course numbers) participate on the project. This makes an LMS gradebook inaccessible. Beyond this restriction, it would be difficult to set up grading for millstones in the LMS because the tasks are individualized and each one is weighted proportional to the work required of the milestone.

Overall, the two metrics providing students consistent feedback throughout the course (*Tri-Weekly Presentations* and *Milestones*) are the second and third most predictive measures – out of the five used – of the final course grade. This result is affirmed by four different analysis methods. Grading rubrics for metrics tri-weekly presentations and milestones are provided in [10] and the appendix. My answer to the first question posed: *are these metrics reasonably predictive of the final grade?* is a modestly confident *yes*, especially for the tri-weekly presentations. But my response to the second question: *can anything be done to help students better understand their status in the course?* is an assured *yes!* Moving forward, I will be sure to conduct a midpoint professionalism review with the design team and provide aggregated feedback to the students.

References

1. G. A. Elamin, N. J. Washuta, J. Righter, and K. Skenes, “Assigning Individualized Grades on a Team Capstone Project,” *ASEE South East Section Meeting*, Marietta, GA, March, 2024. 10.18260/1-2-45507
2. J. K. Estell and J. Hurtig, “Using Rubrics For The Assessment Of Senior Design Projects,” *ASEE Annual Conference and Exhibition*, Chicago, IL, June, 2006. 10.18260/1-2-486
3. L. J. McKenzie, M. S. Trevisan, D. C. Davis, and S. W. Beyerlein, “Capstone Design Courses And Assessment: A National Study,” *ASEE Annual Conference and Exhibition*, Salt Lake City, UT, June, 2004. 10.18260/1-2-13900
4. R. Melnyk and D. Boettner, “An Integrated Approach to Grading a Mechanical Engineering Capstone Design Course at the United States Military Academy,” *ASEE Annual Conference and Exhibition*, Pittsburgh, PA, June, 2008. 10.18260/1-2-3205
5. R. W. Brown, “Autorating: Getting individual marks from team marks and enhancing teamwork,” *Frontiers in Education 25th Annual Conference*, Atlanta, GA, November, 1995. 10.1109/FIE.1995.483140
6. J. M. Hatfield and J. T. Tester, “Assessing Individual Performance Within a Team Using Peer Evaluations,” *ASEE Annual Conference and Exhibition*, Salt Lake City, UT, June, 2004. 10.18260/1-2-13370
7. E. A. Eschenbach, “Using Peer Evaluations for Design Team Effectiveness,” *ASEE Annual Conference and Exhibition*, Milwaukee, WI, June, 1997. 10.18260/1-2-6882
8. D. B. Kaufman, R. M. Felder, and H. Fuller, “Peer Ratings In Cooperative Learning Teams,” *ASEE Annual Conference and Exhibition*, Charlotte, NC, June, 1999. 10.18260/1-2-7880
9. D. M. Dulaski, “Team Grading in Capstone – What the Students Think When They Grade One Another,” *ASEE Annual Conference and Exhibition*, Indianapolis, IN, June, 2014. 10.18260/1-2-23107
10. S. F. Kiefer and T. M. Ericson, “Capstone Design Assessment and Student Motivation,” *ASEE Annual Conference and Exhibition*, Seattle, WA, June, 2015. 10.18260/p.23663

Appendix

Tri-Weekly Presentation Rubric:

	Points Possible	Report 1	Report 2	Report 3	Report 4	Report 5
Quality of notebook entries	3					
Stand-up presentation	2					
Demonstration of tangible progress	5					
Total Points	10					

Milestone Rubric:

Topic (Weight)	**Unacceptable (0)**	**Marginal (7)**	**Exceptional (10)**	**Score**	**Weighted Score**
Research Quality **Weight: 30%**	No published works cited	2 quality **and** relevant works cited	More than 3 quality and relevant works cited		
Analysis **Weight: 30%**	None Provided	Not complete, but correct and useful **or** Complete with minor errors	Everything complete and correct		
Degree of Completion **Weight: 20%**	Work has not been started	Some further quality references needed, does not impede progress of project	Complete		
Confidence **Weight: 20%**	No confidence in analysis or possibility of meeting future plan	Questionable analysis **and/or** Unrealistic possibility of meeting future plan	Nothing questionable in analysis and clear ability to meet future deadlines		
				TOTAL	