2021 ASEE ANNUAL CONFERENCE
Virtual Meeting | July 26–29, 2021 | Pacific Daylight Time
Paper ID #32366

Work in Progress: Examining Engineering Seniors Students' Perception of Justice and Fairness of Grading Practices

Dr. Kaela M. Martin, Embry-Riddle Aeronautical University, Prescott

Kaela Martin is an Associate Professor of Aerospace Engineering at Embry-Riddle Aeronautical University, Prescott Campus. She graduated from Purdue University with a PhD in Aeronautical and Astronautical Engineering and is interested in increasing classroom engagement and student learning.

Todd M. Fernandez, Georgia Institute of Technology

Todd is a lecturer in the Wallace H. Coulter Department of Biomedical Engineering at Georgia Institute of Technology. His research interests are engineering students beliefs about knowledge and education and how those beliefs interact with the engineering education experience.

Prof. Richard Mangum, Embry-Riddle Aeronautical University, Prescott

Richard T. Mangum is an assistant professor at Embry-Riddle Aeronautical University, Prescott Campus. He graduated from Texas Tech University with a PhD in Technical Communication and Rhetoric. He is interested in helping engineering students discover the relevancy of technical communication in their academic pursuits and future career paths.

Work-in-progress: Examining engineering seniors students' perception of justice and fairness of grading practices

This work-in-progress paper reports preliminary results about engineering students' perceptions of the justice and fairness in grading. The paper quantitatively compares students' perceptions between a specifications graded and traditionally graded capstone class and analyzes qualitative comments about students' overall perceptions of grading. Specifications grading, as defined by Nilson (2015), is a system of assigning grades primarily characterized by grading each assignment as pass or no-pass (i.e., fail). Specifications grading responds to the call for grading reform in education (Brookhart, 2011) and builds on grading tenets found in mastery grading or competency-based grading (Bloom, 1971) and standards-based grading (Carberry, Siniawski, Atwood, & Diefes-Dux, 2016; Iamarino, 2014). Specifications grading is proposed as a way to increase students' learning, motivation, and autonomy while also modeling real (engineering) workplace practices. Alongside other techniques like standards based grading, specifications grading is one of several proposed methods of solving issues with grading, assessment, and learning in education (Reeves, 2011). Prior work in engineering education and other fields have demonstrated different implementations of specifications grading for a variety of course topics and designs (Fernandez, Martin, Mangum, & Bell-Huff, 2020). Proponents of specifications grading make many claims about its effect on students' learning and motivation.

Little work, however, has holistically assessed students' perceptions of specifications grading techniques, especially in engineering. Those perceptions are important specifically because the students' experience with new pedagogical innovations may be very different than theorized, especially when techniques deviate from expected educational norms (Deslauriers et al., 2019). Examining how students perceive a pedagogical technique that is significantly different from typical engineering classroom procedures speaks to engineering students' relationship to educational norms in general as well as how the theory-derived claims of specifications grading hold up when implemented in a real classroom. Studying students' perceptions builds on other work about the nuances involved in implementing research-based instructional strategies (RBIS) effectively to provide insight on how new techniques are brought into engineering classrooms (Borrego et al., 2013).

This paper is part of a broader project to understand engineering students' experiences and perceptions with specifications grading, grading in general, and the relationship between grading and learning. Throughout the project, we use justice and fairness as the primary theoretical framework (Colquitt & Rodell, 2015). Justice and fairness provides a theoretically rigorous framework for understanding how students' perceptions, expectations, and experiences interact with new engineering education methods (Martin, Newstetter, & Le Doux, 2019). The specific purpose of this work in progress paper is to study how students perceive grading in a specifications class as fair and just, in comparison to other classes they have previously taken. We do so using qualitative and quantitative techniques.

Study Design

We collected data using a survey from six engineering capstone design sections at a small, private, Southwestern US university. At this university, capstone is a two-semester sequence that is co-taught by engineering and technical communications instructors. Students typically enroll in a separate technical communications (TechCom) course and engineering design course (Capstone) as part of the sequence. Our study population includes 193 students, of whom 65 were enrolled in TechCom and Capstone using specifications grading. The remainder experienced a traditional, points-based, grading approach.

The survey consisted of a 15 item instrument (see appendix), two additional opinion questions, basic demographics, and a single open-ended comment question (optional). We derived the first 13 items from Colquitt and Rodell's (2015) workplace justice and fairness instrument by shifting the focus of each item to an equivalent part of the grading experience. The Colquitt and Rodell survey has been widely

studied and adaptations have been developed for use in multiple contexts (c.f., Rodell, 2013; Colquitt et al., 2019). The original instrument uses four subscales, procedural, distributive, interpersonal, and informational justice. However, Colquitt and Roddell (2015) suggest a two-factor solution that collapses interpersonal and informational into distributed and procedural is also acceptable. In our survey, students responded to the derivative instrument for three different contexts: (1) Courses they had taken in previous semesters, (2) Their capstone course, and (3) Their TechCom course. We plan a more extensive reporting of the instruments' development and validity in future work that is not possible in a WIP.

In addition to the grading justice and fairness instrument, we asked two additional questions in the survey. The first asked students to compare the similarity of grading in their courses to real engineering work, and the second asked students about their preference for the grading system in their Capstone and TechCom courses to their previous course. The similarity question is inherently but purposefully speculative. We are not inspecting the accuracy of students' responses but rather whether they perceive specifications grading as more comparative to the professional work, a claim made by specifications grading proponents (Fernandez, Martin, Mangum, & Bell-Huff, 2020).

Students completed the survey one time, near the end of the fall 2020 semester. After IRB approval, the survey was emailed individually to students enrolled in the six capstone courses. Due to COVID-19 and classroom-size limitations, only the sections taught by the authors were asked verbally to complete the survey, which negatively impacted response rate and biased it towards students in the specifications grading sections. Overall, we received 50 students' responses and removed three students as non-responsive (e.g., responded with all 1s in under a minute) for a final response count of 47 (24%). Of the 47, 21 were enrolled in the specifications grading Capstone and 26 were in traditionally graded Capstone courses. Four of the responses were not enrolled in the technical communications course. We received 22 responses to the optional open-ended comment question.

Quantitative results

As this is a work in progress paper, we focused our reporting of quantitative results on initial comparisons of students' perception of specifications grading. We plan to address evidence of validity for the instrument in detail in later work. Prior to using the instrument for comparisons, we briefly reviewed typical markers of validity. Generally, factor loadings were high, the factor structure aligned with prior work, and measures of EFA model fit were satisfactory.

We carried out two ANOVAs from the quantitative study. The first ANOVA assessed whether the use of specifications grading and course (i.e., Other courses, Capstone, TechCom) affect students' perception of justice and fairness using the instrument average. The second used the additional comparative item to look at students' preference for specifications grading over other courses. In this way, we can separately comment on students' latent perception of justice and fairness from their self-reported preference. Because our sample size is relatively small, our power in these two analyses is somewhat limited, and there is an increased probability of type II error in both ANOVAs. The results of the ANOVA are in Table 1 and a Box plot of the averages by factor level appears in Figure 1.

For both ANOVAs, enrollment in a specifications grading capstone course was a statistically significant predictor. For the first ANOVA, the specifications grading factor was significant but the course students were asked about nor the interaction of course and specifications grading were not significant. This result may suggest that students enrolled in the specifications grading courses simply respond more positively to the instrument. The second ANOVA also shows that the use of specifications grading significantly increased students' preference for the grading in their capstone and technical communications courses over their other courses. As would be expected in a work in progress, more data and deeper explanation are warranted here. However, the results suggest that specifications grading may be well perceived by engineering students.

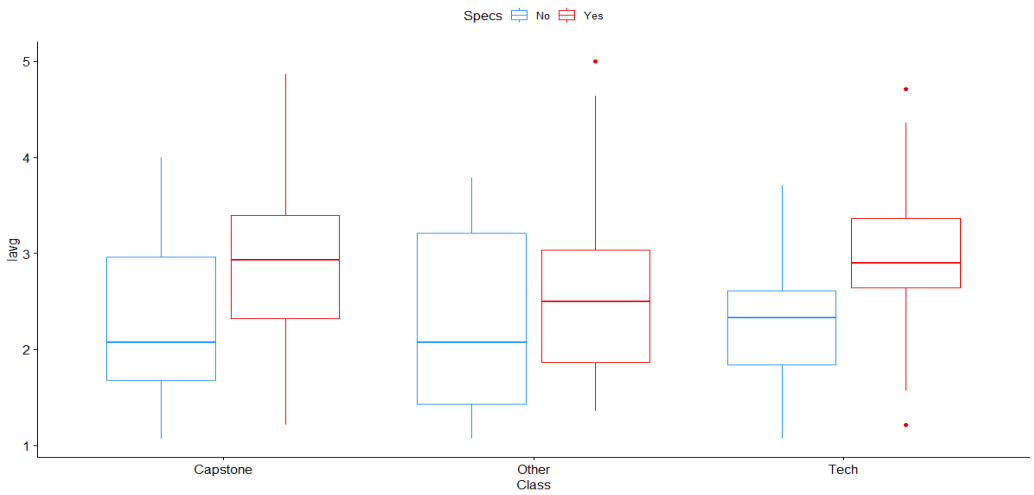

Figure 1 Box plot of instrument overall average score by course and method of grading

Table 1 ANOVAs of instrument scores and grading system preference

ANOVA	**Instrument overall average score**				**Preference over other courses**			
Factor	**df**	**MS**	**F**	**p**	**df**	**MS**	**F**	**p**
Course	2	0.95	1.32	.27	1	0.00	0.00	0.99
Specifications	1	7.56	10.52	**.002**	1	25.85	17.37	**<.001**
Course*Specs	2	0.55	0.77	.47	1	0.07	0.05	0.83
Residual	128	0.72			84	125.07		

Qualitative results

The qualitative results, however, provide a contradictory picture of students' feelings about specifications grading as well as strong feelings about grading more generally. As noted, of 47 responses, 22 students (14 specifications graded, 8 traditional) responded to the optional open-ended question *Do you have any other thoughts you would like to share regarding grading in your courses in general?* Students could respond to that question once at the very end of the survey. Through emergent coding and grouping, we identified several consistent patterns as well as classified each comment as an overall positive or negative perception of grading. Comments ranged from a few words to several paragraphs in length. We plan to use the comments to evaluate whether students' perspectives on grading justice and fairness are fully captured by the current draft of the instrument as well as directions for future work.

The qualitative comments suggest a highly negative perspective on specifications grading. Several comments also express negative perceptions of grading more generally. We identified 10 discrete codes, the full list of which each appear in the appendix with definitions and examples. A typical comment had multiple codes applied (e.g., consistency, team-based grading, the future impacts of grades, perceived grading norms, grading's emotional impact). Overall, we assessed 20 of 22 comments as having a negative perception of grading practices, whether or not specifications grading was used, with only 1 suggested a cohesive positive experience with any form of grading. That comment discussed specifications grading:

"The capstone grading system is very much like what would encounter in industry and I think it is awesome. I do very much prefer it over the way other courses are graded because I feel it reflects my performance better…I do not think the method of assigning letter grades is that difficult to understand, and appreciate there is a clear document which outlines what pathways you have to each letter grade."

In contrast, other comments mentioned aspects of specifications grading, negatively and in ways that indicated students frustration when attempting to map key aspects of specifications grading implementations onto their existing framing of how grading works:

"A clearer grading scheme is needed in terms of points or numbers. What equivalent value in a number or percent grade dictates a 1 or 0 on an assignment? Is it split similar to a rounding scenario? i.e. a 49% or below is a 0 and anything above is a 1? Are the values cut to what the percent values to a grade are? i.e. a 65+% is a D and therefore a 1? Not addressing this allows for a vague grading system that can be weighted by a likeness of the students or project."

Several of the codes and comments suggest that students' experiences with this implementation of specifications grading may not align with the claims made by advocates of the technique, we return to this in the discussion. Other exemplar quotes appear in the appendix alongside the codes and code descriptions.

Several other observations from the comments will influence future work on this project. First, the comments suggest several aspects of justice and fairness that may not be well captured in the current draft of the instrument, such as the influence of teammates or groups on students' grades. Our second observation was that zero comments discussed learning or whether grades accurately reflect student learning. While prior research suggests poor links between grades and learning (e.g., Crowell, 2015), we were intrigued by students' perception of what grades should indicate and the role they play in their lives. For example, multiple comments connected effort and grades such as "I do not believe this accurately portrays the effort level of the team." Students connecting effort and grade outcomes has been reported previously in literature (Singleton-Jackson et. Al., 2010). Students also saw grades, when administered properly, as negotiable. Negotiability is another phenomenon noted in Singleton-Jackson et. al.'s work.

Finally, the comments consistently discussed grading as unfair and a source of explicit, deep, personal grievances directed towards faculty in courses they are currently taking or took years prior. The most explicit example was a student, not in a specifications graded capstone, who commented: *"Most professors grade correctly, others fail me for no reason, play favorites or grade your assignments/exams incorrectly and they refuse to change the grade. I would have already graduated if all professors graded correctly."*

Discussion of our work in progress

The purpose of this study, as outlined, is not to assess whether specs grading is an effective tool for assessing learning. Rather the purpose of this study and our broader work is students' reactions and perceptions to a new pedagogical technique related to grading. The results suggest a number of interesting findings, which we frame through how they will influence future work on this project. Primarily, we are intrigued by the difference (and resulting implications) between the quantitative and qualitative findings given what they suggest about students' reactions to RBIS. The two data sets suggest significant contradictions. As such contradictions do, these contradictions are rather interesting aspects of the phenomenon we wish to observe that provide areas for continued work.

The results of the quantitative analysis suggest positive student perception of specifications grading and a preference for that grading over their prior courses. While further exploration and larger samples are necessary, we are interested in the possibility that experiencing a change in grading system may change the perspective of grading in prior courses. One characteristic of most new RBIS is that the deviations from educational practices students already know are noted and addressed in detail. Here, we intend to further explore students' reactions not just to specifications grading in particular but changes to grading practices in general. This future work includes extensive validity work to expand initial positive indicators on the instrument we trialed here.

The qualitative results, however, suggest the need to better understand how students experience grading and the reliance of that grading on instructor action, unspoken norms, and considerations of what defines a grade. From the perception of the instructors, the specification grading system was extensively explained repeatedly throughout the semester with the differences explicitly compared to traditional grading. However, the students' comments suggest a very different perception on the same course. This difference is not a critique of students, but rather, an observation about the creation of intersubjective space in classrooms (Mortimer & Werscht, 2003; Natahan, Eliam, & Kim, 2007). Students' expression of confusion about specifications grading is broad enough that we theorize that the confusion results in large part from a shift in norms. However, it is equally possible, and an instructor responsibility, that the implementation was a root of confusion. More exploration of this point is needed and in progress, which again links with prior work on grading that such systems are typically left implicit unless deviated from (Gordon and Fay, 2010). Based on these findings, Colquitt and Roddell's (2015) justice and fairness workplace framework appear to be a useful but incomplete framework to explore engineering students' perceptions of grading. Studying perceptions as instructors refine implementation of specifications grading will be an interesting future study.

Beyond specifications grading, the qualitative data also suggest that students feel that grading can often be personally biased, can be inappropriate when linked to teamwork, should reflect effort, and is based on norms which can be done "incorrectly." Whether or not existing grading techniques are effective at measuring learning, students' perception of what a grade should and should not include are useful for instructors to understand.

Since one of the Likert-scale survey questions asked students to compare grading to how they would be evaluated in a job, some students also responded to the open-ended question on how they would be evaluated in industry and describe grading, and specifications grading specifically, as reflecting engineering work in ways that may suggest naïve perceptions of industry jobs. Concepts of how the "real world" is different from school have been shown in the past, and prior work suggests that engineering students often fail to grasp the collaborative nature of the workplace (Dunsmore, Turns, and Yellin, 2011).

Thankfully, students' comments did not suggest one historical role for grading, ranking them and their peers. In any event, the comments suggest that how engineering students experience grading in general is an area in need of significant study. Given the importance that multiple commenters placed on grades, future work here is likely of significant value to understanding students' reactions to RBIS.

More broadly, we do not think these results should scare educators away from specifications grading but rather serve as a reminder of the importance of implementation details with RBIS. The original implementation of specifications grading in fall 2020 involved assignments graded on a 0 (does not meet), 1 (meets), and 2 (exceeds) scale system to address previously experienced concerns of the 'harshness' of a 0-1 only scale. Based on student's reactions, that was changed in spring 2021 to simply 0 (does not meet) and 1 (meets). After this change and other revisions to further explain the grading system and its role, the same students as those enrolled in fall 2020 communicated more positive experiences of specifications grading throughout the course. We plan to continue to use specifications grading. We believe that specifications grading does lead to positive student outcomes and also believe that the nuances of implementing and communicating RBIS matter immensely.

Limitations and Future Work

A major limitation of this study was the survey response rate of 24%. Furthermore, this survey was deployed for the first time at a small private university which is predominately white (62%) and male (73%). As would be expected from a WIP, further data and work is needed. Our qualitative analysis is also limited by the non-random choices to leave or skip the open comments. As is common with end-of-course evaluations, open-ended responses tended to be written by people who felt sufficiently aggrieved

to invest time in commenting. Based on the results, we plan to dive deeper into students' perceptions of grading through focus groups and individual interviews to continue to develop our understanding of these perceptions.

References

Bloom, B. (1971). Mastery learning. In J. H. Block (Ed.), *Mastery learning: Theory and practice* (pp. 47-63). New York, NY: Holt, Rinehart & Winston.

Borrego, M., Cutler, S., Prince, M., Henderson, C., & Froyd, J. E. (2013). Fidelity of implementation of research-based instructional strategies (RBIS) in engineering science courses. *Journal of Engineering Education, 102*(3), 394–425. https://doi.org/10.1002/jee.20020

Brookhart, S. M. (2011). Starting the conversation about grading. *Educational Leadership. Retrieved from https://www.greatschoolspartnership.org/wp-content/uploads/2016/11/Starting-the- Conversation-about-Grading-2.pdf*.

Carberry, A., Siniawski, M., Atwood, S. A., & Diefes-Dux, H. A. (2016, June). Best practices for using standards-based grading in engineering courses. In *Proceedings of the ASEE Annual Conference & Exposition*.

Colquitt, J. A., Sabey, T. B., Rodell, J. B., & Hill, E. T. (2019). Content validation guidelines: Evaluation criteria for definitional correspondence and definitional distinctiveness. *Journal of Applied Psychology, 104*(10), 1243-1265. https://doi.org/10.1037/apl0000406

Colquitt, J. A., & Rodell, J. B. (2015). Measuring justice and fairness. In R. S. Cropanzano & M. L. Ambrose (Eds.), *The Oxford handbook of justice in the workplace* (p. 187–202). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199981410.013.8

Dunsmore, K., Turns, J., & Yellin, J. M. (2011). Looking toward the real world: Student conceptions of engineering. *Journal of Engineering Education, 100*(2), 329-348.

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences, 116*(39), 19251-19257.

Fernandez, T. M., Martin, K. M., Mangum, R. T., & Bell-Huff, C. L. (2020, June). Whose Grade is it Anyway?: Transitioning Engineering Courses to an Evidence-based Specifications Grading System. In *Proceedings of 2020 ASEE Virtual Annual Conference.* https://doi.org/10.18260/1-2--35512

Gordon, M. E., & Fay, C. H. (2010). The effects of grading and teaching practices on students' perceptions of grading fairness. *College Teaching*, *58*(3), 93-98.

Iamarino, D. L. (2014). The benefits of standards-based grading: A critical evaluation of modern grading practices. *Current Issues in Education*, *17*(2).

Martin, C. C., Newstetter, W., & Ledoux, J. M. (2019). Inclusion requires a comprehensive understanding of justice. *Journal of Engineering Education, 108*(4), 453-458. https://doi.org/10.1002/jee.20296

Mortimer E. F. & Wertsch J. V. (2003). The Architecture and Dynamics of Intersubjectivity in Science Classrooms. *Mind, Culture, and Activity*, *10*(3), 230-244. https://doi.org/10.1207/s15327884mca1003_5

Nilson, L. (2015). *Specifications grading: Restoring rigor, motivating students, and saving faculty time.* Stylus Publishing, LLC.

Reeves, D. (2011). *Elements of grading: A guide to effective practice*. Solution Tree Press.

Rodell, J. B. (2013). Finding meaning through volunteering: Why do employees volunteer and what does it mean for their jobs? *Academy of Management Journal, 56*(5), 1274-1294.

Singleton-Jackson, J. A., Jackson, D. L. & Reinhardt, J. (2010). Students as Consumers of Knowledge: Are They Buying What We're Selling? *Innovative Higher Education,* 35, 343–358. https://doi.org/10.1007/s10755-010-9151-y

Appendix

Appendix 1: Survey Instrument

The following questions had a five-point Likert scale (strongly agree to strongly disagree). Past tense was used for prior classes.

1. The grading system is designed with my best interests in mind
2. The grading systems ensures that I received the grade that I deserved
3. There is a clear and fair processes to address any concerns I had with my grades
4. My grades accurately reflect my level of effort
5. The grading system is appropriate for the assignments in this class
6. The grading system prevents favoritism
7. Course assignments have clear instructions about how they would be graded
8. I understand the grading scheme used in this class
9. The grading system gives me sufficient information on why I received a certain grade
10. I can positively influence my grade by doing better quality work
11. My grade accurately reflects my class performance
12. The grading system is applied consistently
13. Assignments are graded in a timely matter
14. My class performance is evaluated similarly to an engineering job*
15. I prefer the grading system in this class to the ones in my other courses* [capstone/technical communications course only]

** Items 14 and 15 were not derived from the Colquitt and Rodell (2015) survey and instead serve as a comparison to other grading system references students might rely on – such as comparison to professional engineering roles or other courses.*

Appendix 2: Codes, descriptions, and representative quotes from qualitative comments about grading

Coded pattern ***Description***	**Representative quotes(s)**
Communication *Comments about how information on grading is made available to students*	"The grading scheme was explained once at the begging*(sic)* of the semester in under 5 minutes and never mentioned again." "I do not think the method of assigning letter grades is that difficult to understand, and appreciate there is a clear document which outlines what pathways you have to each letter grade."*
Consistency *Comments about how students perceive grades to vary (or not) in unexpected ways within the same system*	"Capstone grading seems based on how much the professor likes you and your team that day and on your teams perception of you, it has absolutely nothing to do with the work you produce."
Team based grades *Comments discussing the tension between team based work and grades being assigned to individuals that may not be entirely based on the individual*	"I feel as if my grade is MORE impacted by my team members because there is only so much work that I can commit to on a project document." "It's a bit annoying to get the same presentation grade as the rest of my group. Most of them are quite terrible presenters and make everyone fail the presentations. I believe the presentations should have one group grade for the content of the presentation and one individual grade for speaking and presenting the information."
Future impact *Comments that note or reference future implications of grades and the effect such implications should have*	"I understand why this grading scheme is in place form an Engineering point of view, but from a student who has a GPA to manage for grad school, it's terrible" "I'm an undergraduate and my grades are the most important aspect of my life."
Grading norms *Comments that assert or imply the existence of 'correct' processes for assigning grades*	"Most professors grade correctly, others fail me for no reason, play favorites or grade your assignments/exams incorrectly and they refuse to change the grade. I would have already graduated if all professors graded correctly."
Emotional impact *Comments that discuss emotions intertwined with grades and grading*	"We are not given instructions or an example and then we are given very harsh feedback, and in the end we still get pretty good grades. It is very confusing and stressful to deal with as a student." "it seriously destroys my mental state and also negatively influences morale"
Grading beyond learning assessment *Comments that discuss factors beyond course learning outcomes that should influence grades and work evaluation*	"The grading caused my team to speed*(sic)* more time trying to meet every single point for a document because our grade hinges on every assignment instead of actually working to improve our design. There wasnt any time to work on the design when everything else took so much time." "I do not believe this accurately portrays the effort level of the team"
Timeliness *Comments about timeliness of grading and feedback*	"In regards to my capstone I believe some things were graded weeks ago but they have not appeared where I should be able to see them in Canvas. I have no idea what my actual grade is yet."
Agency *Comments that identify unexpected grade outcomes or suggest an inability to affect the outcome of a grade.*	"In regards to previous classes the implemented systems had little effect on my enjoyment of the class. It was the teachers that chose to manipulate those systems to increase or lower the difficulty of their class. For example: I loved taking chemistry (my least favorite subject). That professor graded everything fairly and allowed anyone and everyone to argue back a few points on tests and homework."
Industry comparison *Comments comparing academic experience against (perceived) understanding of industry work*	"The capstone grading system is very much like what would encounter in industry and I think it is awesome."* "The grading system is absolutely terrible and nothing like 'industry'."
Notes: Quotes selected for representativeness. Instructors noted several comments made objectively false statements about the availability of rubrics, syllabi, and course information. Analysis did not consider the course the comment came from although such information was at times identifiable. * This quote comes from the only solely positive comment on grading in our qualitative results.	