Peer Advising at Scale: Content and Context of a Learner-Owned Course Evaluation System

Alex Duncan
Georgia Institute of Technology
Atlanta, GA
Alex.Duncan@gatech.edu

David Joyner
Georgia Institute of Technology
Atlanta, GA
David.Joyner@gatech.edu

ABSTRACT
Peer advising in education, which involves students providing fellow students with course advice, can be important in online student communities and can provide insights into potential course improvements. We examine reviews from a course review web site for online graduate programs. We develop a coding scheme to analyze the free text portion of the reviews and integrate those findings with students' quantitative ratings of each course's overall score, difficulty, and workload. While reviews focus on subjective evaluation of courses, students also provide feedback for instructors, personal context, advice for other students, and objective course descriptions. Additionally, the average review varies by course overall score, difficulty, and workload. Our research examines the importance of student communities in online education and peer advising at scale.

Author Keywords
Peer advising; online communities; graduate education; CS education

ACM Classification Keywords
• Applied computing~Collaborative learning • Applied computing~Distance learning

INTRODUCTION
Academic advising at scale in online programs can pose various challenges: students who are uninformed about university policies; a low advisor-to-student ratio; and rapidly changing courses. Peer advising – in which students are advised by other students – can assist with academic advising at scale. It may be formally instituted by a university [1], which has been shown to be effective [2, 3], and it can strengthen a student community, leading to improved student performance and retention [4].

Our work looks at a course review web site for massive (over 10,000 total students in Spring 2019) online graduate programs in computer science, data analytics, and cybersecurity. The site allows students to write reviews of courses in addition to providing an overall score and ratings of difficulty and workload. It currently has almost 3,000 reviews total for ~30 courses.

In this paper, we develop and apply an original coding scheme to the free text of the reviews and use this to summarize review content in conjunction with an analysis of the quantitative aspects of the reviews. The purpose of our research is to: examine what information students choose to share with their peers; learn how this information relates to quantitative assessments of courses; and understand how the components of course reviews contribute to the peer advising community.

METHODOLOGY
To develop our coding scheme for the free text portion of the reviews, we looked at 1,373 reviews (the number of valid reviews available to us at the time). 75 reviews were selected and coded at the sentence level to develop a grounded coding scheme. The final scheme contained six codes: Advice, Review Context, Course Description, Evaluation, Feedback, and Other. The code definitions are shown in Table 1. 52% of the remaining reviews (comprising 677 reviews, or 6,746 sentences) were then coded to conduct the analysis for this paper. Both intra- and inter-coder reliability were assessed and found to be within acceptable limits.

RESULTS
Figure 1 shows the code breakdown as a percentage of the total number of sentences. Note that longer reviews will have a stronger influence on these percentages. Figure 2 shows the average review makeup, generated by determining the percentage of sentences with each code in each individual review and then averaging these percentages. This summary removes the bias of longer reviews. Figure 3 shows the percentage of reviews where a code appears at least once. This figure addresses the bias that could arise if some types of information require more sentences to communicate than do other types.

In addition to writing a free text review, students also rate a course based on an overall score (from 1 to 5, with 1 being "strongly disliked" and 5 being "loved"), difficulty (from 1 to 5, with 1 being "very easy" and 5 being "very hard"), and workload (in hours per week). Tables 2, 3, and 4 show the average review makeup for all courses as broken down by

L@S '19, June 24–25, 2019, Chicago, IL, USA

ACM ISBN 978-1-4503-6804-9/19/06.
https://doi.org/10.1145/3330430.3333660

Definition	Examples
Advice	
Recommendations involving prerequisite knowledge, courses to take before or in conjunction with a course, or the best way to progress through a course were included in this category. Also included were warnings or reassurances about taking a course and information about future offerings of a course. Advice is particularly characterized by being targeted at the reader.	*"Knowledge of python is a must."* *"Keep up to pace with lectures, projects, and reading material."* *"If you complete all the assignments and take the extra credit test, you should get a good grade."*
Review Context	
This category included statements about the reviewers themselves, such as their coding experience. Facts that were specific to one semester of a course were also included, along with non-advice, non-evaluative statements about how the reviewer or other students progressed through a course.	*"Because I took this in the summer, the workload was much higher than you will see in other reviews."* *"I have no formal CS background."*
Course Description	
This category contained statements that provide objective information about a course, such as the number of projects or average final grade in a course. Any factual statements about a specific semester of a course that are likely generalizable across many other semesters are also included.	*"The class has 3 assignments, 3 projects, and a midterm and final."* *"All of the coding is done in Java, using the IntelliJ and Android IDEs."*
Evaluation	
Statements in this category were subjective and related to the reviewer's opinion of a course, often involving likes or dislikes. Statements about a reviewer's dislikes were only included in this category if they weren't actionable; otherwise, they were grouped into the Feedback category.	*"The concepts presented in the lectures aren't too difficult to wrap your mind around, and I found many of them very interesting."* *"This class can get a bit boring."*
Feedback	
Actionable statements regarding aspects of a course that a reviewer disliked or wanted changed were grouped into this category. These statements were broadly applicable and not specific to one student. Although most of these statements were usually recommendations for course changes, some of them were about aspects of a course that were beneficial and should not be changed.	*"Since the projects are fairly trivial, it seems silly to spend so much time on design and discussion, so hopefully in the future they use projects that are a bit more complex, or maybe having to deal with changing requirements or incomplete requirements."* *"I'd suggest slowing down some of the more critical lectures or providing more examples."*
Other	
Statements that didn't fall into any of the other 5 categories were grouped into this category. These were often post-semester outcomes, musings about education as a whole, or fragments that only make sense in the context of previous sentences.	*"Also, because of the practical side of this class, I was able to get a job as a Junior Data Scientist!!"* *"In real life requirements tend to be vague and it takes a good analyst to fill in the blanks."*

Table 1. Coding scheme definitions and examples

overall score, difficulty, and workload respectively. Average review makeup was used for this analysis to standardize the "units" of information being studied, since course ratings are given only at a review level. Students may report any number for workload, so we defined the categories seen in Table 4 as if workload were rated on a Likert scale to facilitate our analysis. This means that not every category had an equal number of reviews; the ">40" category comprised only 3 reviews and is ignored in our discussion of trends.

DISCUSSION

Although statements coded as Evaluation appear more often than any other type of statement, statements coded as Advice, Course Description, and Feedback appear at least once in the majority of reviews, and statements coded as Review Context appear at least once in 45.9% of reviews.

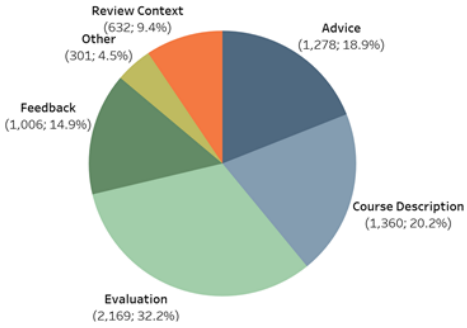

Figure 1. Content of reviews based on total code usage.

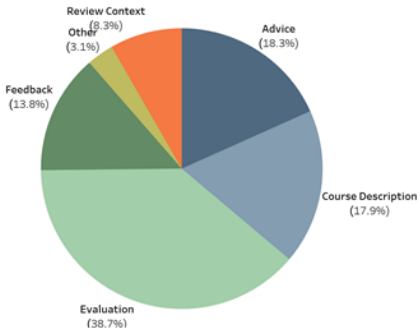

Figure 2. Content of reviews based on average review makeup.

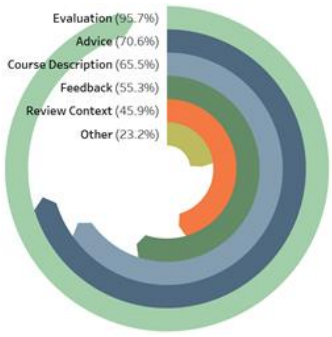

Figure 3. Content of reviews based on percentages of reviews with at least one example of each code.

Non-evaluative aspects of reviews are perhaps easily communicated with just a few sentences, whereas evaluative aspects require more sentences to describe. The appearance of multiple categories of information within individual reviews, in addition to the different average review makeups for courses with different ratings of overall score, difficulty, and workload, suggest that the various categories of statements work together to form a more wholistic review.

As expected, there is a general decrease in the amount of feedback in the average review as course overall score improves. Interestingly, the average review increases in the number of evaluative statements with higher overall scores. Presumably these types of statements become more positive with higher overall scores, implying that students are more inclined to share positive opinions about courses than negative ones. There is also an increase in percentage of the average review devoted to advice as overall scores. It is possible that, while students still share advice about courses with lower overall scores, this advice is more succinct than it is in reviews of courses with higher overall scores, where students may encourage enrollment in the course in addition to sharing their thoughts on how to succeed in the course.

As may be expected, advice peaks for very hard and very easy courses, suggesting that students want to encourage each other to enroll in very easy classes and discourage each other from enrolling in difficult classes. Difficult classes also probably warrant more guidance on how best to prepare for and proceed through the class.

With increasing workload, there is a general increase in the percent of the average review dedicated to contextual information; in contrast, there seems to be no distinct trend for review context related to either course overall score or difficulty. This suggests students typically use contextual statements to qualify the amount of time they invested in coursework, particularly when they invested a lot of time.

Overall Score	Advice (%)	Review Context (%)	Course Description (%)	Evaluation (%)	Feedback (%)	Other (%)
1	11.53	6.20	14.29	29.79	32.50	5.71
2	16.61	6.57	14.77	35.33	22.46	4.27
3	15.78	9.49	18.73	36.30	16.70	3.01
4	18.14	7.36	20.30	40.29	11.59	2.32
5	21.11	9.82	16.64	40.48	8.96	3.00

Table 2. Average review makeup per course overall score.

Difficulty	Advice (%)	Review Context (%)	Course Description (%)	Evaluation (%)	Feedback (%)	Other (%)
1	22.69	10.30	14.95	37.49	11.91	2.66
2	16.13	6.37	20.67	41.46	12.62	2.75
3	17.85	7.24	18.34	39.33	14.03	3.21
4	15.99	10.49	17.87	37.62	15.17	2.85
5	25.91	9.59	13.37	34.88	12.63	3.62

Table 3. Average review makeup per course difficulty rating.

Workload (hours/week)	Advice (%)	Review Context (%)	Course Description (%)	Evaluation (%)	Feedback (%)	Other (%)
0-10	17.02	6.94	19.04	40.70	13.67	2.63
11-20	18.80	8.55	17.41	39.07	13.48	2.69
21-30	19.53	11.44	16.62	33.82	13.97	4.63
31-40	24.72	10.57	15.46	22.18	19.94	7.15
>40	19.40	16.15	8.29	29.75	12.37	14.04

Table 4. Average review makeup per course workload rating.

The average review for a course with a weekly workload of 0-10 hours contains 40.70% evaluation. This percentage plummets to 22.18% for courses with a weekly workload of 31-40 hours. A similar decrease in course description percentage is evident. Seemingly counterintuitive, these decreases coincide with a general increase in Advice and Feedback. Instead of complaining about a high workload or listing the course assessments, perhaps students desire to help their peers by alerting them to the workload, advising them about how to tackle it, or providing recommendations to instructors about how to make it more manageable.

This inverse relationship between evaluation and feedback can be seen across course overall scores, difficulties, and workloads; it may relate to the difficulty of distinguishing Feedback from Evaluation. For example, a statement that the course requires too much reading may be the reviewer's opinion, or it may be feedback to the instructor to reduce the amount of reading. Perhaps Feedback and Evaluation combined represent a single category, and the observed differences are the result of a shift in tone or language.

CONCLUSION

Applying an original coding scheme to reviews on a course review web site for online graduate programs, we found that students often provide subjective evaluations of courses and contextualize these evaluations with course descriptions and personal details. The information a student provides in a review varies with how the student rates the course in terms of an overall score, difficulty, and workload.

The web site we examined is specific to technical graduate programs, so our findings may not generalize to less technical fields or programs with different student demographics. Future work may focus on performing similar analyses on course reviews in these other contexts.

Additionally, more actionable insights may be obtained by making our coding scheme more granular. This would allow us to explore subcategories of information, such as what types of advice students provide, and in what contexts.

We could also further examine how average review makeup changes based on multiple course ratings. How do students discuss easy courses with high workloads? What about courses students enjoyed but found extremely difficult?

Lastly, this course review site forms only one part of the peer advising community for these programs. Future research could look at the role it plays in this community and examine interactions in the entire community to provide a broader understanding of peer advising at scale.

REFERENCES

1. Barman, C. R., & Benson, P. A. (1981). Peer advising: A working model. *NACADA Journal*, *1*(2), 33-40.
2. Brown, C. R., & Myers, R. (1975). Student vs. faculty curriculum advising. *Journal of College Student Personnel*.
3. Murry, J. P. (1972). The Comparative Effectiveness of Student-to-Student and Faculty Advising Programs. *Journal of College Student Personnel*, *13*(6), 562-566.
4. Thomas, L., Herbert, J., & Teras, M. (2014). A sense of belonging to enhance participation, success and retention in online programs.