The Future of Engineering Education
2024 Annual Conference & Exposition
Oregon Convention Center
Portland, OR . June 23 - 26, 2024
Paper ID #41447

Using Oral Exams to Assess Individual Contributions on Team Projects

Dr. Matt Gordon, University of Denver

Dr. Matt Gordon is Professor of the Department of Mechanical and Materials Engineering. His research areas include numerical and experimental plasma physics, chemical and physical vapor deposition, electronic packaging, and bio-medical engineer

Irvin R Jones, University of Denver

Using Oral Exams to Assess Individual Contributions on Team Projects

Abstract
Most, if not all, mechanical engineering programs recognize the value in teamwork. To satisfy ABET requirements, a senior design course is required, and teams are almost always involved. Often, teams are involved in other prior courses as well. One difficult aspect, however, with teams, is ensuring that all members contribute satisfactorily/equally. It can be all too common that one or more members of a team do the majority of the work and other team members do less work.

At the University of Denver (DU), the co-authors taught a junior-level design course during the winter and spring quarters of the 2022-23 academic year. To address this possible issue of ensuring that all teammates contribute, we introduced an individual oral exam held in the last 2 weeks of the course. The class had 61 students, and each oral exam was scheduled for 15 minutes. Both instructors participated in all exams. Although conducting these individual oral exams required just over 15 hours, we believe it was a sound investment of our time. After the exams, it was very clear who contributed and who did not.

Introduction

Teamwork is a key part of any engineering curriculum to better prepare students for life after academia. Any academic program with an external advisory board will no doubt be told this from their industry partners. And, of course, ABET directly mandates it [1]. ABET now actually defines 'Team' – "A team consists of more than one person working toward a common goal and should include individuals of diverse backgrounds, skills, or perspectives." And from Student Outcome 5, students must have "an ability to function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives." Thus, all programs are clearly instructed to include, and assess, teamwork.

While it is easy to put students in teams, it is much more difficult to assess individual performance and ensure that all contribute as desired. The most common course to require teams is senior design. ABET again mandates this course saying that the curriculum must contain "a culminating major engineering design experience that 1) incorporates appropriate engineering standards and multiple constraints, and 2) is based on the knowledge and skills acquired in earlier course work [1]."

As a result, there have been many studies on this topic – how to effectively assess individuals on teams. In broad terms, instructors must decide on one of two general approaches. On one extreme, once a team is assigned, instructors can let each team plan and manage the process for completing the project, meaning it is acceptable however the team decides to get the project done. Typically, this approach leads to every individual on a team getting the same score. This approach is easiest for the instructor since it really doesn't matter what an individual contributes. All that matters is that the project was completed. However, this approach can create a lot of stress for the team, especially if they have different objectives. One student might be happy with a 'C' (a passing grade) while another wants an 'A'. The one desiring the 'A' often then must carry the team and resents the fact that the other members obtained a high mark without putting in the work. On the other extreme, each member of the team must materially participate in all aspects of the project (e.g. a team can't decide one person will write the report, another will do the calculations, etc.). The instructor's challenge in this paradigm is to then determine how each person performed individually, independent of how well the team performed. Most instructors attempt something in between these two extremes.

As anyone who has taught a class with team projects knows, such courses are typically more difficult to teach than more traditional courses. If the project is open ended, as most are, inevitably teams will develop issues unrelated to engineering topics, and most instructors are not well equipped to handle these. Sometimes one team member will see the instructor to report that some team members are not contributing to their satisfaction. Another team member might complain that other team members are not respectful of them personally or of their contributions. Sometimes instructors hear nothing, but later find out that one student was given virtually nothing to do because another teammate felt that they couldn't be trusted with any important tasks. Regular team meetings with the instructors, or team surveys such as CATME [2], can help uncover these issues, but such problems are not always caught. And even if they are caught, it isn't always clear what the best resolution is.

Marin-Garcia and Lloret wanted to assess teams to avoid shirking, among other things, and tried using group process assessment by the instructor – assessing how the group functions and not just the results the group produces [3]. Their approach focused on in-class instructor evaluations of student participation by means of an observation grid. They concluded that it was possible to minimize shirking and implement such a method even with a large number of students. However, they concluded there were definitely limitations, one of which being that this method used class-time which might otherwise be used for other (better?) purposes.

Andreu-Andres et al. focused on assessing both critical thinking and participation of team members in open ended projects [4]. They relied on students self-assessing after carefully guiding them with focused exercises. They concluded that this method could work but wondered if it would scale to large groups.

The use of oral presentations is common in team projects, but predominantly their use in assessment is for presentation skills (Meseguer-Duenas et al. for example [5]). Presentation skills are another required outcome in ABET [1].

For our purposes, we wanted to assess individual knowledge/competence/contribution rather than presentation skills and decided to use both an assessment by the students through CATME and an individual oral exam given in the presence of both instructors.

The University of Denver has regularly used CATME in both this junior level Integration course as well as Senior Design. As far as we know, no one at our university has used an individual oral exam for this purpose. We do implement such an exam of course at the graduate level – qualifying exams, thesis/dissertation defenses, but nothing at the undergraduate level. Presumably, most/all would agree that such an oral exam would be effective, but also would be very time-consuming.

The Course

At the University of Denver, we do have a senior design course where students are indeed put in diverse, multi-disciplinary teams, but for this paper, we are using a junior-level, two quarter long, course (ENGR 2610/2620), where students are also put in teams to complete an open-ended design project. Somewhat unique to us, both design sequences enroll all our engineering majors (Computer, Electrical, and Mechanical engineering) for a truly inter-disciplinary experience. Figure 1 shows the inter-disciplinary structure of ENGR 2610/2620.

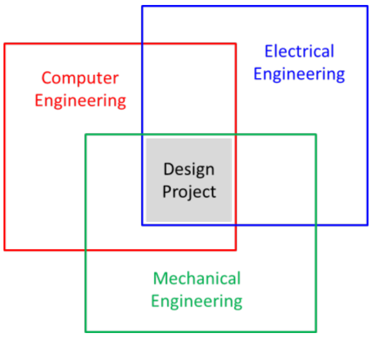

Figure 1. Inter-disciplinary structure of ENGR 2610/2620

ENGR 2610 (Integration I) is offered in the winter quarter followed by ENGR 2620 (Integration II) offered in the spring quarter. Both courses are ten weeks in duration. The two co-authors were assigned as instructors for the course – both teaching it at DU for the first time. One of the co-authors (Dr. Jones) had taught a similar course at another university. In part, because we knew in past offerings of this course that not every student contributed to all areas, we decided to assign individual lab work during the first 8 weeks of the winter course so that all students had experience with the key systems design and development skills needed for the spring project. The lab experiences focus on the areas of overlap between the engineering disciplines (Figure 1). Thus, all students completed labs, individually, involving microcontrollers, microcontroller interfaces, control algorithms and programs, power (mechanical and electrical), motors, servos, sensors and actuators, CAD software, and mechanical construction and design. An additional benefit of the experiential learning associated with the individual lab assignments was that students learned how their discipline relates to other disciplines in an inter-disciplinary environment and how to apply their skills to satisfy the functional needs of an inter-disciplinary system.

For the last two weeks of the winter quarter, we put them in their teams (teams were put together by the instructors using input from the students – they could name one or two persons they would like to team with and one or two persons they would NOT like to team with – and balanced by our three engineering majors – Computer, Electrical, and Mechanical engineering). Our objective was teams of four, but with 61 students, we opted for one team of five. As a team, they then completed a two-week lab which was more open-ended. During the spring quarter, they remained in their teams and were assigned their project.

Each team needed to design and build an autonomous robot that would complete three tasks. For the first task, their robot needed to pick up a box that would either be large or small, and red or blue. Based on which box they had, their robot would deliver it to the correct location (there were 4 locations – one for each possibility). For the second task, their robot would need to navigate a simple maze made up of a random set of walls with each wall having a single opening on the left, right, or center. The robot had to move from one end of the course to the other, while avoiding contact with the walls and staying within the boundaries of the course. For the third task, they simply needed to determine the maximum weight they could push or pull, and then also carry, up an incline.

The Oral Exam

Each of the 61 students were asked to sign up for one 15-minute time slot to complete their oral exam. We created 64 slots (16 hours) to accommodate this, spread out during the last two weeks of the 10 week spring quarter. All of the slots were during their normal lab times so as not to conflict with other commitments. During week 10, student teams also had to sign up for testing times to demonstrate that their robots could perform the three required tasks. There was some overlap in those times (performance demonstration and oral exams) but between the two instructors and the four teaching assistants, we could ensure that at least two instructors and/or teaching assistants were present at each. Sheets were posted on the professor's door so that students could sign up anytime and see what was available. For those signing up towards the end, we were prepared to open up other times if needed, but it was not. The time slots were back-to-back with no time in between, and we used time blocks of either 1 or 2 hours. Doing four straight exams was pretty easy, but it was more taxing during the two-hour blocks (8 straight exams). Two students failed to show up at their allotted times. One had a good excuse and we were able to reschedule to an open slot. The other student just never responded to our outreach and never took the exam.

Both instructors were present during every exam. We each had the same sheet with the questions and space for our comments. We each would give an overall score at the end – we would of course do this right after the exam and before the next one started. Scores were given as follows:

0 – No answer or very poor answer
1 – Response was vague lacking foundation/support
2 – Response was missing many details; some important points were not addressed
3 – Response covered several of the important points with supporting details
4 – All or most of the important points were covered with supporting details.

In ENGR 2610/2620, inter-disciplinary teams work on an inter-disciplinary project. In this scenario, it was imperative that every member of a team participate in the design, development, and testing of their system. The oral exam given to each student would test the breadth, depth and scope of knowledge each student had about their project. The exam would cover areas of general system design and behavior, individual contribution, and team dynamics (i.e. individual challenges of working in an inter-disciplinary team, and challenges the team faced).

The oral exam had 6 questions, many with multiple parts. Instructors would alternate asking questions, and either instructor could follow up on any item as desired. We quickly fell into a good rhythm where we generally each asked students the same questions.

The 6 questions were as follows and presented to the students orally and only during the exam:

1. Briefly tell us about your contribution in the area you led?
2(a). Explain the drive train of your system and how it works.
2(b). Where was the center of rotation of your system?
3. Why did you choose the wheel configuration you chose? If you had to do it over, would you stick with that configuration?
4. What was the greatest challenge you faced as an individual? As a team? How were they resolved?
5. Describe the algorithm used to accomplish task 1. Same for task 3.
6. How does your device determine box size and color? What other options were considered? Knowing what you know now, do you think you have the best solution?

We believed, correctly it turned out, that the first question would allow each student to feel more comfortable to start the exam. For all questions, we expected different levels of responses depending on what they focused on and what their major was. For example, we expected more from mechanical engineering majors for questions 2 and 3, and more from electrical and computer engineering majors for questions 5 and 6. However, we wanted all students to be somewhat knowledgeable of the entire team effort. Question 4 proved to be the most interesting for the instructors even if it did not typically have a strong influence on the final score.

As needed/desired, students could use a white board and markers to answer these questions. Probably 90% of the students did write/draw on the board to support their answers.

Results and Discussion

As mentioned, both instructors were present for the 60 oral exams, each lasting 15 minutes. One of the 61 students never showed. 33 scored in the 3-4 range, 25 scored in the 2-3 range, 2 scored in the 0-1 range. These scores reflected the averaging of the two instructor's scores (each instructor would decide on the overall score based on their sub-scores), but there was very good agreement between them.

For the CATME scores, we used their adjustment factor (without self). From the CATME web page [6]:

> "The adjustment factor compares an individual student's ratings with the average ratings of everyone in the team. This helps to see if the student was harsher than the average, or less harsh. There are two different adjustment factors, "Adj Factor w/Self", which includes the student's self-rating in the calculation, and "Adj Factor w/o Self", which does not. These adjustment factors can be used to "score" the student's ratings."

Using this score, 55 students scored between 0.94 and 1.08. These all were considered good. Of the remaining 5, 4 scored between 0.78 and 0.88 which were considered low. And one student scored a 0.59 which was considered extremely low.

We believe both sets of data are interesting and note that it was much more time efficient to collect the CATME data (students completed that on their own for a required homework score). Again, it required about 15 hours of each instructor to collect the oral exam data.
What was most interesting was that no student scored low on both. The students with the two lowest scores on the oral exam each obtained a CATME factor of 1. Thus, it appears while they appeared to know very little about how their team's final design worked, they were considered 'good' teammates. Conversely, the students with the 5 lowest CATME scores all did well in the oral exam. It bears consideration of how the composition of the team affected the CATME evaluations. Most all teams were composed of members that wanted to work with each other. It is likely that students would be reticent in critically reviewing their teammate.

Thus, while we conclude that both scores are useful, they measure different attributes, both of which are likely important. For our purposes, a low score on either resulted in a lower grade for the student as we indeed felt both were important. However, we initially assumed there would be a better correlation between those two scores.

After the class was completed and grades submitted, one of the instructors informally asked students from that class what they thought of that oral exam. The consensus was that there was perhaps the most value in the students' knowing about the exam. Most of the students said that that alone caused them to pay more attention to what their teammates were doing than they thought they would have done otherwise. They also said they were pretty nervous about the exam (just them with both instructors and unknown questions – they were NOT given the questions ahead of time) but all said it went fine. They said questions were reasonable and no one reported feeling unduly uncomfortable during the exam.

Conclusions

In an attempt to better determine an individual's contribution on a team project, we implemented a 15-minute individual oral exam. While this was a heavy investment of the instructors' time, we believe that just having the students know there would this exam led to better performance, and the results of the exam clearly showed where two students had not understood how their team's design functions. We also collected teamwork data from CATME and were surprised to see little correlation with the exam results. No student earned a low score in both. We conclude that the CATME score better reflects the non-technical side of being a teammate while the exam focused on the technical knowledge. Both are important, but these are two very different attributes.

In short, the exam was considered a success and worth the time and will be implemented again during the 2023-24 academic offering of this course.

Referenes

[1] "Criteria for Accrediting Engineering Programs, 2024-2025." [Online]. Available: https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2024-2025/. [Accessed April 16, 2024]

[2] MW Ohland, HR Pomeranz and HW Feinstein. "The Comprehensive Assessment of Team Member Effectiveness (CATME): A New Peer Evaluation Instrument," in *Proceedings of the 2006 ASEE Annual Conference*, Chicago, IL, USA, June 2006

[3] JA Marin-Garcia and J Lloret, “Improving Teamwork with University Engineering Students. The Effect of an Assessment Method to Prevent Shirking,” WSEAS Transactions on Advances in Engineering Education, V5-1, 2008

[4] MA Andreau-Andres, M Garcia-Casas, and B Rising, “Assessment of Student Participation and Critical Thinking in Engineering Students’ Teamwork,” WSEAS Transactions on Advances in Engineering Education, V6-12, 2009

[5] JM Meseguer-Duenas, A Vidaurre, J Molina-Mateo, J Riera, and RM Sala, “Validation of Student Peer Assessment of Effective Oral Communication in Engineering Degrees,” IEEE Revista Iberoamericana de Tecnolgial del Aprendizaje, V13-1, 2018

[6] “What is the adjustment factor?” [Online]. Available: https://info.catme.org/instructor-faq/what-is-the-adjustment-factor/. [Accessed April 16, 2024]