Peer marking on a Group-based Problem Solving

Faculty: Natural Sciences

Department: Physics

Module name: Comprehensives

Module leader: Dr Mitesh Patel

Format: Group work

Level: Undergraduate Year 3

Approximate number of students: Full cohort (About 250 students) split into two terms

Duration: Ten weeks (One term)

Module ECTS: 15

Assessment overview

Group-based problem solving is a continuously assessed component of the core year 3 undergraduate physics third year Comprehensives module – a module designed to augment and consolidate students' problem-solving abilities using material covered in the first two years of the core undergraduate degree. In group-based problem solving students work in teams of 25 to solve a complex problem. The assessment runs over an entire term, in which the students are expected to organise themselves, and, ultimately, present their findings in a half-hour seminar. Unlike most other assessments, their mark is not provided by the judgement of the module lead, rather, the entire assessment is peer-based, with the other teams grading each group's seminar. Each individual mark is then determined by peer-to-peer marking within a group, based on the overall grade of the group.

Design decisions

Rationale for the design

The initial motivation for a group-based project was to improve the core physics degree's provision on training regarding collaborative skills of the students. There is a perception amongst some that Imperial sometimes exhibits rather competitive environments, which may lead to students developing relatively strong individual academic skills but relatively weak teamworking abilities. The idea of the group-based problem solving is hence to provide an assessment that is as close to the real world working environments as possible, where collaborative groupwork is the norm. Initiated during Curriculum Review, the project ran for the first time for the 2021-22 year 3 cohort.

In real-world working environments, groups tend to be externally evaluated mainly on their final product, whereas internal team dynamics has minor importance from the point of view of the external observer. In this assessment teamwork is given as much credit as other higher visibility components, such as, for example, scientific rigor. The peer-marked component is designed to reflect the environment in a professional scientific research group environment such as a research committee.

Each team is eventually awarded one overall mark in a final seminar day at the end of the project. This mark is composed out of the marks given by all the other teams based on their final work. A second contribution to the individual marks comes from the peer-wise assessment throughout the duration of the project. Each team member assesses the others twice and provides extensive feedback. Hence, the individual team members weight gets accounted for, but not as strongly as it might be in conventional forms of group projects.

The structure of the assessment involves a ten-week period in which the teams work self-organized on a large-scale problem. During this time, they organise their own structure. The teams rotate the chair of their first meetings while eventually electing a spokesperson. The spokesperson

Anatomy of Assessment

Peer marking on a Group-based Problem Solving

is meant to have rather a coordinator-ship rather than a leadership role. Additionally, a dispute resolution procedure has to be developed. While they work mainly independently, there are non-assessed weekly tutorials where the module lead mainly provides academic support with the physics involved. As the groups are expected to collaborate, the module lead should only be involved into the resolution of conflicts if the dispute resolution procedure fails to settle the conflict.

Eventually, each team has to present their work in a 30-minute seminar and then answer questions. In 2022, there were five groups per term. The four non-presenting groups then mark the seminar, which accounts for the majority of the mark. The seminar tries to imitate an environment similar to a grant proposal, and uses a somewhat sophisticated marking scheme. The teams give marks across five categories, with constraints in the number of marks available. This means that the scores are limited, with each mark being only available for distribution a set number of times. For example, a team can only award two “Outstanding”, while “Good” can be given four times and so on. This forces the teams to be thoughtful of the marks given and to eventually give the correct number of marks across all categories and presentations. Staff do also provide marks, which are used as a moderation tool. As in 2021-22, the student’s marks entirely aligned with the marks given by staff, thereby requiring no moderation by staff.

Alignment with learning outcomes

As the assessment name indicates, group-based problem solving tries to enhance collaborative skills and improve students’ abilities at working in a team. The group will have to be able to present their findings in a seminar, even if some group members did not collaborate as effectively as others. Hence a great emphasis on teamwork and conflict resolution is made, together with the need for each team to come up with their own organisational structure.

As this assessment is part of the Comprehensives module in the year 3 curriculum, it is also meant to test the students ability to apply their knowledge and skills acquired over their first two years of studies. As none of the students are expected to have prior knowledge of the material used in the problem assigned, they have to use all their problem-solving abilities to collaboratively tackle the problem at hand.

Since the assessment runs over an entire term, it is also very important for the students to continuously work on the project and have a working organisational structure at hand, both for each student individually and as a whole group. They will have to delegate, carry out tasks that connect various areas of physics and do general research. Hence critical thinking, combinatory abilities and problem-solving skills are key in this assessment additionally to teamwork and collaborative skills.

Practicalities

Assessment and provision of feedback

Group-based problem solving is assessed entirely on a peer-based system. As mentioned before, there are two components to the individual marks given: the peer assessment during term, and the mark given to the whole team based on their presentation. This second mark is then shifted based on the mark an individual student achieved in their peer assessment during the project.

The peer-based mark consists of single peer feedback in week 3, contributing to 20% towards the final mark, and a second feedback in week 10 carrying the majority of the weighting with 80%. The intention is that individuals use

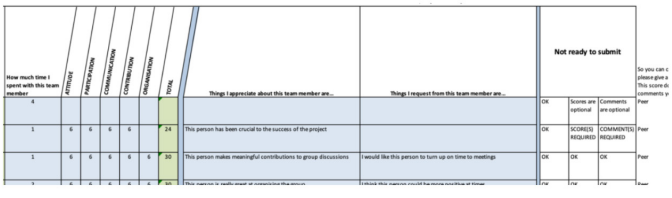

Figure 1:Sample of part of the form that allows students to give each other feedback and mark their respective contribution during the peer-assessment process. Given marks are weighted by the time spent collaborating and feedback is encouraged.

Anatomy of Assessment

Peer marking on a Group-based Problem Solving

the early feedback to reflect and adjust their working methods and styles how best they see fit to improve, both personally, and for the benefit of the team. Individuals will find out, for example, if their peers' perception is that they are not contributing enough, or if they are doing something particularly constructive. In week 3, this can be used as either a confidence boost if someone is performing well, or as a point of realisation if someone needs to contribute more.

The feedback and marks are weighted by the amount of time two individuals collaborate. This way, the mark & feedback given by a direct & frequent collaborator has a greater impact than the mark & feedback by an occasional acquaintance within the group. There are four different categories used to quantise the amount of collaboration between two specific students:

- I was assigned to work with this person on the same task. To complete the task successfully and fairly, I needed to communicate with this person a lot (multiple times a week or for long periods of time).
- I was assigned to work on a similar area of the problem with this person. To align our goals, I needed to communicate with this person regularly (once or twice a week).
- We didn't work on the same part of this project, but I attended regular meetings with this person (once a week or once every two weeks).
- I've not really interacted with this person.

Each student needs to mark at least five students they have interacted with. The marking is done across five different categories, viz (i) attitude, (ii) participation, (iii) communication, (iv) contribution and (v) organisation (see the marking scheme later in figure 2). The students are also able to provide written feedback, while answering the prompts "Things I appreciate about this team member are…" and "Things I request from this team member are…". These written feedbacks have shown to be highly effective, as students in 2021-22 provided constructive and detailed feedback about the performance of their team. There is also a space for feedback to be given to the team as a whole. Students can also give themselves scores and comments, as a chance for self-reflection.

The seminars form the conclusion to the assessment with each team presenting their work in a 30-minute seminar, followed by ten minutes of questions. Afterwards, the other teams are given 10 minutes to provide provisional marks for the presentation that has just occurred. The marks are provisional as the teams are given time at the end of the day to recalibrate and adjust all grades having seen all presentations. Marks are given across five equally weighted categories, viz. (i) creativity, (ii) scientific rigour, (iii) achievement, (iv) teamwork, and, (v) presentation quality. The redistribution of marks after hearing all presentations is of special importance due to a sophisticated system of distributing marks. Instead of each group being entirely free in deciding the mark of a presentation, there are constraints of how many marks each team can give in total. Hence, a group can only give a limited number of "outstanding", or "satisfying" scores. This can be seen as an implicit way of ranking all teams across the different categories, similar to how it is done in grant proposals. The system is put in place to achieve a friendlier environment and to simplify the decision process of achieving a ranking of the presentations observed. Since a team will have observed four teams and will mark them in five categories, they will award 20 scores at the end of the day.

Calculating the final mark

This assessment exhibits a somewhat sophisticated marking scheme. There are several components contributing to the final marks given to the students. Not only is each grade composed out of the groups mark together with a weighting factor arising from the feedback given by one's group, but also the initial components to the final grade arise not trivially.

In method for calculating group contributions is shown below in figure 2:

Anatomy of Assessment

Peer marking on a Group-based Problem Solving

Person a gives person b mark m_{ab} with weighting w_{ab}. The weighting is determined by the distance levels 1, 2, 3, or 4. Level 4 always has $w_{ab} = 0$. Person a gives N_a people a score (from only level 1,2,3, also ignoring self-assessment scores). There are N students in the team.

Assessor a gives an average score of,

$$\bar{m}_a = \frac{1}{N_a}\sum_i^{N_a} m_{ai} \tag{1}$$

This average is used to normalise all the scores given by assessor a,

$$m_{ab}/\bar{m}_a \tag{2}$$

which tells us if each person, according to Assessor a, is performing above or below average in this category. This step allows us to account for certain markers being more generous or stingy than others. Each peer is then awarded the weighted average of their assessor-normalised scores,

$$s_b = \frac{\sum_j^{n_b} w_{jb}\,(m_{jb}/\bar{m}_j)}{\sum_j^{n_b} w_{jb}} \tag{3}$$

where n_b people have given marks to person b (in levels 1,2,3). The next step is to then normalise this to the average score of the team,

$$\bar{s} = \frac{1}{N}\sum_k^N s_k = \frac{1}{N}\sum_k^N \frac{\sum_j^{n_k} w_{jk}\,(m_{jk}/\bar{m}_j)}{\sum_j^{n_k} w_{jk}} \tag{4}$$

so that we can find who, within the team, performed above or below average. The scores are normalised to the team's average,

$$s_b/\bar{s} \tag{5}$$

Finally, to find the redistribution of the grades between team members, these scores are now converted to shifts,

$$\frac{s_b}{\bar{s}} - 1 \tag{6}$$

The shifts from the two rounds are simply added together with the appropriate weighting,

$$s_{b,\text{final}} = 0.2 \times \left(\frac{s_{b,1}}{\bar{s}_1} - 1\right) + 0.8 \times \left(\frac{s_{b,2}}{\bar{s}_2} - 1\right) \tag{7}$$

The shifts are then transformed into a percentage shift by multiply by a factor f, so that the final percentage shift for one round of peer marking is

$$f \times s_{b,\text{final}} \tag{8}$$

The factor in 2021-2022 is $f = 10$.

As a final consideration, any penalties the students have accumulated are also added here. No submission of the peer-feedback form is equivalent to -2.5% shift, and non-attendance at the final seminar day is -8%.

Figure 2: The mathematical procedure used for calculating students' marks following peer assessment

There are a limited number of points available to be awarded at the final seminar day, which is done in a style that imitates research proposal environments. To calculate a score from the day, the rankings are converted into numbers: Outstanding = 5, Excellent = 4, Good = 3, OK = 2, and Satisfactory = 1. These are then averaged together to give a numerical mark for each team, ni. The algorithm used to compute the score is show in figure 3:

To convert this number into a percentage, we find the mean $\bar{n}$ and standard deviation σ_n of the numerical scores for the whole day which is used to convert the numerical mark into a percentage,

$$p_i = \bar{p} + \left(\frac{n_i - \bar{n}}{\sigma_n}\right) \times \sigma \tag{9}$$

where $\bar{p} = 65\%$ and $\sigma = 8\%$ are the mean and standard deviation of the target distribution.

The final grade is calculated by,

$$\text{final grade} = \text{seminar score} + \text{peer feedback shift} \tag{10}$$

Figure 3: The Mathematical procedure used for calculating the final grade

The peer-assessment component is weighted by the time spent collaborating between two students, and both the groups mark and a student's weighting factor then need to be combined following an appropriate grade distribution. This design process was devised prior to first implementing the assessment in the academic year 2021-22.

The assessment has ran twice 2021-22, with half of the year three cohort in the first term and the second half completing the assessment in the second term of the academic year. Between both executions, staff had the chance to reflect on the assessment and introduce slight improvements allowing a smoother running of the assessment the second time.

Student's perspective

Students are expected to work 75 hours each on this assessment. Since there are 25 people in a team, the project they have to tackle exhibits a great complexity. The main takeaway for students was the ability to work in large groups, since this assessment forms the first encounter of students with groups as large as 25 people. Yet this also posed the greatest challenge, as students report that they felt a disproportionate amount of time was spent trying to organise themselves. As working environments offer an initial skeleton organisation structure which is not in place in this group project, students felt that it would be very beneficial to have slightly more support at coming up with an initial structure of organisation. Additionally, the nature of having large groups lead to some students being too concerned with the organisational aspects of the project, at the expense of devoting time & energy to contributing to the physics, which was handled by other members of their group. This aspect was not appreciated as much by students, and was also something flagged as an unanticipated problem by the module lead. Yet, even though this assessment has mainly been experienced as challenging, students tended to have a strong and positive appreciation of the idea and philosophy of this assessment.

Further observations of assessment type

- For the 2021-22, it was been observed that teams hardly collaborated among each other. One suggested cause of this is that the groups find themselves in competition from the outset, even though the whole exercise is never posited as a competition.

Interviewees: Dr Mitesh Patel and Dr Isabel Rabey

Anatomy of Assessment

Peer marking on a Group-based Problem Solving

- Another approach which has been found to be detrimental to the teams by the project lead is the splitting of teams into coding and non-coding subgroups. Even though there is a great deal of complex physics involved, the projects are all (in the current format) largely coding projects. One suggestion to improve this is that all students should contribute to the coding, at least in some small part.
- A group of 25 is a large group to manage. While teams of this size are common in workplace, they would always either be deliberately organised into, or would naturally form subgroups. Asking students to organise groups themselves would probably not by that reflective of the workplace as those structures would be put into place for the employees linked to their job titles and roles. An alternative strategy is to create this structure for the students and then ask them to volunteer for different subgroups and roles. There are clearly several options for how to organise a group of this size should module leads wish to pursue this.

Advice for students

- It is important to find an organisational structure and put this in place very early on in the assessment.
- This assessment emphasises teamwork and collaborative skills and should be approached as such. If the group dynamic fails, teams had a hard time being able to reach their conclusions.
- Students are equipped with approximately 3 hours of introductory material, which students are able to refer back to. This gives a lot of advice on how to approach this assessment and the physics involved.
- This assessment really offers the opportunity to make use of each and everyone's individual strengths. There might be some students who have a great sense of seeing the bigger picture and are hence very suited for more organisational tasks. Some other students might have an immediate grasp of the physics at hand or have a good sense of explaining the subject and hence would do best in tasks related to the coding or presentation. A team would benefit greatly from exploiting these individual strengths.

Interviewees: Dr Mitesh Patel and Dr Isabel Rabey