Automated Mechanical Engineering problem sets

Faculty: Engineering

Department: Mechanical Engineering

Module name: MECH50010 Fluid Mechanics 2

Level: Year 1 and 2

Format: Homework sheets and Progress Tests, online

Approximate number of students: 200

Delivery mode: online; both remote and on-campus

Weighting and credit: summative assessment is 5% of the module (low stakes)

Module ECTS: 5

Module Type: Core

Assessment overview

These assessments focus on automated feedback provided for mathematical, short-answer questions to problem sets for undergraduates. The assessment feedback is both formative and summative.

The self-study exercises consist of homework sheets set frequently throughout the autumn and spring term with accompanied feedback. The summative assessment consists of 'progress tests' worth 5% of the 5 ECTS module (these are "low stakes" assessments), set at the end of the first term. There are 17 formatively assessed homework sheets in total across the module, which are set weekly in autumn and spring term. The formative self study exercise questions include a range of types, from basic practice, to reinforcement, to extended challenges. They are not designed to 'test' understanding, but rather to give students the necessary exercise to learn.

Design decisions

Rationale for automated marking

Student homework sheets were not marked in the past on this module, so automation allows feedback on formative assessments where previously there was none. More specifically, it automates 'low level' feedback that stops students getting stuck on questions, and allows them to go deeper into their homework before requiring help. The subsequent contact time with teachers is higher quality due to deeper discussions. See this article for more.

Automation for summative assessments improves staff time management at a busy time of year and increases the speed and consistency of feedback to students.

The reason for the low stakes assessment is two-fold; it is an entry point to the technology for the teacher (allowing development of the system), and it introduces students to this type of assessment and feedback without a high mark penalty if they get things wrong. The progress tests are a good vehicle for this innovation as they are summative (so students take them seriously), but the weighting is low.

Rationale for the design of the Lambda software

Initially the assessments were hosted on Mobius platform. However, the following issues were identified:

- General user experience (student and staff)
- Restriction on programming language to program feedback (Maple only).
- Restriction on error-carried forward capability.
- Restriction on feedback capabilities - because it is 'grading' focussed.
- Restriction on analytics.

The Lambda feedback software was designed to meet the needs that Mobius couldn't. In addition to that the software will soon allow the

teacher to parameterise a question, so that the exact content seen by a student is unique to them and so the correct answer is different for each student. There are obvious issues with the integrity of remotely-delivered summative tests, so this system allows teachers to change some (or none) of the questions, which prevents students colluding on their answers.

Questions design

For the self-study exercises, the same considerations were made as for any worksheet-style assessment, including, from the student perspective: the purpose of the assessment e.g. which Intended Learning Outcome is being assessed (what concept, what skill, what knowledge, what misconception?); how long will it take to complete?; How clear are the instructions for the assessment? How much interest/engagement does it generate for the student? How easy is it to navigate through the assessment, e.g. how good is the user experience?

From the staff perspective, the teacher needs to consider how to construct a question to enable generation of the formative automated feedback, so must consider the structure of the question, and any potential ambiguity in the answer, e.g. use of different symbols for mathematical concepts.

Fit with other assessments and the programme/module

The formative feedback assessment is new and innovative; the feedback provided does not cause any known disruption to any other module, and supports students' learning of essential concepts in a way that is beneficial to the higher level modules that they take in later years in their degrees.

For the summative assessments, changes (such as from paper based to computer based and from manually marked to automatically marked) were applied across the whole cohort to ensure consistency of the marking.

Practicalities

Preparing students for assessment

As preparation for the formative assessments students are given general advice on study methods and feedback literacy. Feedback literacy is a two-way thing between students and teachers: students need to receive a lot of feedback before they get skilled at knowing how to react to it; teachers need to provide quality feedback. Students can check their understanding of the module content and concepts during bi-weekly tutorials with staff.

As preparation for the summative assessments students are given a briefing in a lecture and are given written information about it on Blackboard; they are directed towards the ILOs, the syllabus and the module descriptor. They also take a readiness test (to check they can use the software). For one cohort (who were particularly stressed) a mock test was run too. Students are encouraged to complete the formative worksheets as preparation for the summative assessment.

Monitoring student progress

As all the assessments are online, useful analytics can be easily generated to track student engagement, e.g. showing peaks in online activity on the day of the lecture; showing that many students access the homework sheet as soon as it is set, then engagement tails off; that there is a gradual decline in student engagement through the time, but that overall students do keep working and keep trying to catch up.

Above: temporal access to the system in one module of 208 students. Lectures were on days 18 and 25. This data is from the last two weeks of term.

Feedback arrangements

Feedback for both formative and summative assessments is automated , as discussed above in the explanation of the Lambda feedback software. It is completely consistent and objective, but also manually checked. Any changes that need to be made (e.g. if an error is spotted in a question) are applied algorithmically to the whole cohort. Students can ask questions about their feedback during tutorials throughout term, or in a dedicated tutorial feedback session in January, after the summative assessment has been completed.

For the summative assessment the feedback is not instant, as teachers check the grading before releasing marks. Feedback is currently limited to grades (marks) with no comments.

Interviewee: Peter B Johnson
Role: Principal Teaching Fellow

Online adaptations

The summative tests were designed to be delivered online, and can be taken remotely (they are currently taken in-person, in computer rooms, but worked well entirely remotely during the pandemic). The online mode of assessment hasn't affected the design of the assessment, except that it increases the importance of parameterised questions (which are bespoke for each student) for the summative test.

Advantages of the assessment type

- Formative feedback is entirely automated. It is consistent, objective and is manually checked.
- If any changes to the marking are required these can be applied algorithmically to the whole cohort, which saves time for the marker.
- If a student thinks that their formative feedback is incorrect, they can "flag" it on the system, which will alert the module co-ordinator, who can then go in and apply any changes if necessary.
- Formative feedback is a huge bonus for the students, as they didn't get any in the past. Students are more motivated to learn and seem happier with the feedback from the module.
- It is easy to monitor student engagement, and potentially to intervene when a student is not engaging (e.g. by alerting their personal tutor).
- The advantages to students of the automated feedback are: timeliness (the formative feedback is instant, and they can react to it and continue through their homework without getting stuck); richness and personalisation of feedback (the students answer difference questions, so receive feedback that is bespoke to them); consistency of feedback due to automation.
- Advantages to teachers of the automated feedback are: higher quality contact time, as less time is spent talking students through small mistakes in the homework sheets; insightful analytics, e.g. level of engagement and competency with the questions; improved student experience and enjoyment of the module.
- Integrating assessment that is formative and developmental in nature but counts summatively for credit towards the degree can be a good way to encourage early, and sustained student engagement.

Limitations of the assessment type

- Summative assessment is still hosted on Mobius and for both students and teachers the user experience is poor, for example, there were a lot of browser issues with TRAs in Mobius. In terms of how the window is laid out there is a lot of poorly-used space on screen, and students have to scroll down pages to see content.
- It would be best if staff used the Mobius system directly, but they are not all trained in best practice and some are concerned about the level of support they will receive if they need help designed their assessment questions.
- Inaccurate feedback is a big problem when the marking algorithms incorrectly evaluate a student expression. For the summative assessments, once the results and answers are released students who think that their question may have been marked incorrectly are incentivised to get in touch with the module co-ordinator to query their result, because if there is an error in the feedback programming, this might result in their mark being increased. For the formative assessments, students are less likely to get in touch if they think the feedback is wrong (though the new "flag" system seeks to redress this, see advantages section).

Advice for implementation

- Formative: it takes time to prepare content, and to refine the feedback.
- Summative: more workload upfront; new constraints on question development (restricted by what can be automated). To be most effective it requires thinking differently about the types of questions that are set for assessment, and that can be a difficult adjustment process for the teacher.
- From employability perspective, think of ways how the development of transferable skills can be supported through preparation for the exam through encouraging group revision and highlighting how group revision can support development of interpersonal skills, negotiation skills and time management skills.
- It is important to ensure that software used allows for changes to be made to the layout of the question and the exam to make it accessible to all students. In terms of the font there is often an assumption that Times New Roman is a good

font to use while in reality it is really difficult for anybody with specific learning difficulties to process. In terms of the layout of the exam questions on the page, having to scroll down between the question and an answer can be challenging hence the question and answer should be visible together without the need to scroll. Another consideration needs to be given to where the buttons are placed and avoiding placing 'next' and 'submit' buttons close together as students with visual perceptual difficulties might find this challenging and accidentally click the wrong button. If a screen reader is required it is important to make sure that the text is accessible. Consider presenting multiple choice options with greater spacing between them especially if answers are very similar visually.

- It is important to consider how the questions are displayed to make the output inclusive. It is best to present multiple choice options with greater spacing between them especially if answers are very similar visually.

- Discuss your software choices with your Faculty EdTech team

- When introducing low stakes assessments with a formative function it is important to consider whether the attached credit doesn't take away from the formative focus, i.e. the focus on learning. It is important to consider the overall assessment burden for the staff as well for the students, i.e. can it be marked within the allocated timeframe, can appropriate feedback be provided so that support can be put in place? All of these considerations should be given when designing assessments of this kind.

- Integrating assessment that is formative and developmental in nature but counts summatively for credit towards the degree can be a good way to encourage early, and sustained student engagement. The disadvantage is that student can perceive this as an extra, continuous pressure as these tests also 'count'. It's important to regularly reinforce that they are small weighted and that completion of them that is more important than the mark received.

- Allowing some time in the curriculum to help students develop feedback literacy will greatly support students with their uptake of feedback and as a result will help them become a much more proactive and independent learners. Feedback literacy can be developed through open discussion around what feedback is and how it benefits students' learning, allowing opportunities for self reflection around performance and feedback, openly discussing emotions around feedback and how to process comments to benefit learning and finally, designing assessments so that there are opportunities to apply feedback to future assignments. It is best if education around feedback literacy starts early on in the degree so that there is sufficient time for practice and so that the skills that students develop can be applied throughout the programme.

- In order to make exams inclusive allowances should be made for students with declared learning difficulties. If the purpose of the tests is monitoring progress then adding extra 25% can be overkill, however if the assessment feeds into the final mark in any way it should attract extra time. If a test is conducted in person this adjustment can be easily made by allowing students stay longer in the examiner hall. If assessment is automated it needs to be reprogrammed to allow extra time for specific individuals. Sometimes staff mind find themselves in a situation where the technology does not allow to adjust the time for specific students in which case every one should be given more time to complete. There have been many studies that found that students finish within the initially allocated time limit hence the only students that benefit from extra time are the ones who needed the adjustment in the first place.