Electronic Mid-term Exams

Faculty: Engineering

Department: Electrical and Electronic Engineering

Module name: Assessments across different modules in EEE with specific references to Communications module

Degree: Across EEE programmes

Level: Across Y1 and Y2 undergraduate courses

Approximate number of students: 100-200

Duration: approx. 50 minutes

Weighting and credit: 10%-20% of the final grade (usually 10% in Year 1, 20% in Year 2

Module ECTS: Variable

Assessment overview

Mid-term exams are designed as low stake assessments with more of a formative function delivered in the middle of the module (usually 6 weeks into the term) with the main aim of checking students' progress and offering appropriate support for future learning. The exams usually consist of between 1 and 15 questions and are delivered and marked electronically via Wiseflow software.

Design decisions

Rationale for the mid-term exam

Previously a lot of modules in EEE were relatively assessment heavy, mostly with high stake end of term exams. Mid-term exams were introduced to offer more distributed, interim assessment points allowing the lecturers and students to formally monitor learning and progress more closely.

> 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 - have they got time to mark everything and get it back so it's properly formative and give good feedback to students so that if you're using it to help students to check their progress and offer an appropriate support, that has to have a very quick turn around time. Electronic marking can certainly help with that.

The mid-term exams allowed departments to distribute lower risk 'smaller' assessments more evenly, 20-30% of module assessment is lab assessment done throughout the term and 10-20% is the mid-term exam and end of term assessment accounts for 50-60%, depending on the module and the programme. This has spread the assessment load for both staff and students, encouraged student engagement with the material throughout the term (as opposed to the previous tendency to shift study to final exam revision) and provided mid-term feedback to both staff and students, while there is still time to learn and improve for end of term assessment. Removing the pressure around the end of term exam also aimed at reducing the stress associated with high stakes exams. Having an exam that accounts for 100% of students' mark created a lot of pressure. Now students approach the end term exam knowing that if they do well in the mid-term and the lab assessment they are approaching the end of term exam with 40% of the course completed. This is intended to remove the pressure, however, whether it truly does remains to be seen as especially in Year 1 students do not tend to distinguish between low and high stake assessments and do not fully appreciate what this means in terms of preparation.

> 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

Insights colour key

Educational Developer

Inclusivity

Learning Designer

Registry

Careers

AI

Interviewee: Zohaib Akhtar, Communications module lead, the digital lead for the department and personal tutor*
***The Case study is written up from the perspective of the role of the digital exam lead for the Department but with references to specific examples from his experience as a module lead and personal tutor.**

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. Watch this video on when it is not advisable to make reasonable adjustments.

From pedagogical perspective a decision has to be made around whether the aim is to spread the assessment load or unload with such a design and why formative assessment should carry a mark. In some ways having a single stressful load is easier to manage than having multiple loads causing a similar level of stress to staff and students. Hence such an assessment design doesn't necessary unload and assessment load but rather spreads it over a period of time potentially causing same level of stress to students who can't distinguish between low and high stakes assessment. There are a couple of reasons why a formative could carry a mark: better student engagement, because you specifically want to capture whether the students understood something that has been done in the early part of the term because you need to know that in order to know how and why to build on it for the latter part of the term, or it could be to practice stressful exams. It is important to consider those decisions in light of your module. In either of these cases an important consideration is if assessment design allows for such return of feedback that would make a difference to student learning.
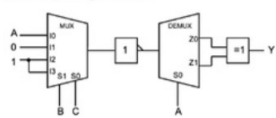
Figure 1

Rationale for electronic delivery
Electronic delivery has long been a part of EEE assessment. As the number of assessments increased, so did the marking and feedback requirements. In order for those assessments to fulfil the role that the teams wanted them to play, i.e. monitor progress and offer feedback that can support learning, those needed to be marked fast. Electronic marking and feedback helped to facilitate that.

Questions design
The type of questions that are asked is partially dictated by the software that is currently used and partially by the nature of the learning being tested. Most tutors were initially drawn to multiple choice questions (MCQs), perhaps because of assumed software limitations, but this is not a preferred question type for Engineering where most end of term exams test analysis and problem solving. It is possible to achieve this on Wiseflow (the software used) and design questions that test these skills using, for example, 'fill in the blank' type of questions (numerical), free text answers for mathematical expressions, fill in the table etc. See Figure 1 and 2.

Figure 2

One thing to note is that we do not ask questions that require complex mathematical expressions as answers even though it is possible to do that with the software. The reason is that it would require significant training of students to type this on the software. Students are very much used to typing those complex equations on a piece of paper, then to some extent are able to transfer it to a Word document but might not have the same ability to do it on LaTex (that is integrated into the software) especially in the context of assessment that is so high stake. Hence if students are tested on those

complex maths equations, the preference is to do it on a piece of paper and upload a photo of the answer for it to be marked manually or the question is written in the form of an MCQ as in Figure 3.

Figure 3

A lot of consideration has been given how to use the available software to best serve the needs of the students and the best types of questions to assess. Software constrains are an important thing to consider when designing exam questions. An even more important consideration is whether not testing those complex mathematical equations is an intended part of the module. If yes then the software might be inappropriate or some solutions need to be found to allow for testing what needs to be tested. The pedagogy should be the driver not the technology.

When moving exams online or when designing exams for online environments. Some questions can easily translate online, some will need a level of tweaking and some will need a complete reconstruction and a different way of asking a question or changing the format to more of an open book one. What is important to know that the knowledge and skills can still be tested effectively but might require a different way of thinking hence will require some effort.

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

module tend to focus on testing knowledge and the theory part of the course (more on theory in the Fit with other assessments section). Understanding basic engineering concepts is important for Year 1 students and the mid term exam focus around assessing this basic understanding that then can be expanded on throughput the year to build the foundations for testing higher order thinking skills in the end of term exam.

When writing questions it is also important to give information how answers should be given, for example if units are needed etc as this could result in marking correct answers as incorrect and vice versa.

Having done the assessment for several years allowed us to develop question banks which means we have a good amount of questions available that can be reused every year without the fear of students having access to past answers.

AI is very good at solving knowledge based questions testing facts or direct relationships as well as these questions that do not require a lot of diagrams, figures or graphs as it can't read those well at this point. This means that it can successfully solve about 30% of first year problems and 20% of second year problems. This score is relatively low because the questions were already designed more in the style of open book exams. Questions that are testing advanced knowledge are safe as currently they are too complex for AI to solve. They can be solved by breaking them down and explaining them to the AI, which is what a more expert user would be able to do, however if a student is able to break the problems down, they would also be able to solve them without the help of AI so the need to use the software becomes redundant.

AI software can be used to help generate exam or more so pop quizzes questions. Those can be multiple choice questions or Fill-In-The-Blank questions. Questions can be generated based on learning outcomes or objectives. The command could look like this: "Generate a multiple-choice question with exactly 4 options based on the learning objective: 'Introduction to Python Programming'. Output in CSV format of the following structure: question, correct option (one of the four options), option1, option2, option3, option4. This query would be executed as many times as needed (e.g., 30), gradually coming up with numerous potential questions. The teacher who

...requested these would then sift through the results and select the ones that were useful (ie. were relevant, meaningful, and made sense), thus creating a pool to pull questions from while randomly generating pop-quizzes for students.

One thing to note that the questions that tend to be suggested by some software (for example ChatGPT) are more aligned with the US based curriculum.

Automated answers design

When using automatically marked exams one of the important things to consider is the amount of differentiation within an allowed answer and ensuring that all allowed variations are entered into the system. This is not something that you need to do in the same way on a paper test but is extremely important in a computer test. Therefore ensuring that you take into account and input every correct answer that the computer would mark as incorrect is a big task.

To ensure that the questions and answers work, we test each question on GTAs who are asked to do those questions as students. This allows us to see if any types of mistakes come up with certain answers hence allows for better quality assurance of the exam.

Rationale for the software used

The software used to manage the electronic midterm exams is WISEflow. There were two main reasons why this particular software was found most suitable: flexibility in the type of question that can be used and real time progress monitoring (which is unique for this type of software).

The College offers formal support for a number of platforms that can host online exams. There are a lot of things to consider when choosing an appropriate platform for hosting digital assessments, for example whether the grading is as important as authoring, whether negative marking needs to be included, whether you want to embed pulling or banking and how feedback is supposed to be returned. Starting conversations early with your learning technology team will help you take all of the important choices into consideration and choose the platform that serves yours and your students' needs best.

Previously, the Easter test and the Christmas tests that were also electronic were conducted on Blackboard. The only possible question type was mainly MCQs. One big benefit of WISEflow it allows a lot of question types. This means that 90% of typical exam questions in the department can be translated to an electronic format. The software also allows us to customise questions for students through reshuffling the order in which they appear as well as amending the values that students work with, which helps to stop collusion.

Another big feature is the ability to monitor student progress in real time. This means that throughout the duration of the exam you can see how much time students are taking on each question, what percentage finished the test at what point in time. The software does not monitor any other activity (i.e. what is appearing on students' screens or doesn't use their camera to record what they are doing physically) but having the data around how far into the test students are allows us to support students who might be having technical issues or helps to identify any questions that the cohort is generally struggling with. We use that information in the following years to adjust the assessments.

When BYOD (Bring Your Own Device) in person mid-term exams were conducted post pandemic, the software also allowed to use the lockdown feature, which means that when students start the test their screens get locked – they can't exit the test and they cannot open anything else. They can only do it once they are given a password from the invigilators. This

also allows us to monitor when students submit their answers. Generally we do not want them to submit before the allocated time. This is to stop any communication between different test groups. Real time progress monitoring also provides evidence as to students' claims around connection issues (as the software records any connection breakdowns with time stamps) if the midterm is conducted remotely. Students also can be contacted by the invigilators if the invigilator notices that a test hasn't been started etc. This allows early fixing of issues.

In terms of equipment requirements and how software was used, currently everything is conducted via a browser – students just open the browser and do the test there. Prior to COVID when the tests were conducted on site and an installation of software that would block students' screens so that they couldn't access anything else was required. Currently those tests are conducted at home hence there is no need for this kind of software as students are in an environment where they have access to multiple devices. This, however, required rethinking some of the questions to make them less googleable.

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 or eyesight problems 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.

Fit with other assessments and the programme/ module

The mid-term exam is an additional assessment method introduced as a diagnostic/ formative tool but carrying sufficient amount of credit to ensure students approach it with diligence. It is designed to monitor progress on the theoretical aspects of the module that provide underpinning for the more practical aspects assessed in the labs. There generally is no overlap between what is assessed in one assessment type but rather the basic concepts assessed via mid-terms provide the base to develop other skills assessed later on.

Using Communications module as an example and the ILOs below:
1. Model noise in wireless communication systems
2. Understand the performance measures for analogue and digital communication systems, and compare the performance of different modulation techniques
3. Explain the fundamental limits information theory sets on the performance of communication systems;
4. Calculate the entropy of a random variable, and design the Huffman code for a given information source
5. Design, implement and characterise hardware/ software implementation using the NI USRP software defined radio platform
6. Understand the role physical layer plays within larger communication networks

We have a lab assessment in the form of a lab report and orals but in the theory assessment we are trying to focus more on aspects of the module that are not overlapping with the lab assessment so that the students are assessed on a range of knowledge and skills covered within the module. For theory module a mid-term exam is the most appropriate option to test theory and its application. What theoretical means in this context is different to the textbook definition of theoretical. For us a theoretical questions can be related to analysis or design, but whatever they are,

When deciding on your approach to AI consider the purpose of the assessment or the questions. The aim of introducing mid term exams was more formative in nature, I.e. to monitor progress and better distribute students' learning. The small stakes nature of the assessment as well as question design that aligns more with open book exams means that the 'threat' posed by AI could be minimal. If exams contribute to a large proportion of the grade, and are therefore high stakes, you may consider holding those assessments in a room on campus where internet access will be disabled and therefore the use of AI will be close to impossible. This strategy will be dependant on class size and available infrastructure. Another strategy would be using an online invigilation tool.

they are essentially maths questions. You write down a few equations, you try to solve them, then you give the answers. So when we say theoretical these are questions that ask for example how would you describe the trend when X changes? What happens to Y? This is further assessed at a deeper level through an end of module exam hence there is a close relationship between mid-terms and final exams, i.e. students have an opportunity to practice on the types of questions that could be to an extent replicated in the higher stake end of term exam.

As mid-term exams have been introduced across the majority of year groups, there is an element of continuation as the students' progress through their studies. This means that the effort put into preparing students to approach them offers longer term benefits and a level of security and familiarity when it comes to an assessment type.

> When planning a module level assessment diet it is important to be clear which module level ILOs are addressed by which assessments. This is going to help reduce overall assessment burden. So a mapping activity can be useful.

Practicalities

Preparing students for assessment
While it is safe to say that the majority of EEE students are computer literate we still do a lot of mock tests (one mock test before every mid-term) to help them get used to the software and practice types of questions they might be asked. The first mock is mandatory to ensure that every student has some practice. If a student does not take the mandatory mock test, they are not allowed to take the mid-term. The mandatory mock test uses past year's questions. Throughout the year students are also offered to attend additional mock tests whose purpose is getting them used to the software and the types of questions and answers that they might encounter. Those type of questions are not related to the content of the module but are rather general. The reason why real questions are used only once is to not to exhaust the bank that will be use for the actual assessment.

Marking arrangements
The marking is set to be done automatically by the software. There are, however, some limitations to the question types that can't be automatically marked and this boils down to descriptive types of questions. These questions are rare for us to ask at the end of term exam. If those questions do appear on the mid-term, they need to be marked manually by the tutor. These types of questions can however be redesigned in the form of for example MCQs or other types allowed by the software.

The software checks for the equivalents of correct answers for mathematical expressions but those need to be prepared in advance and put into the allowed answers box. Unfortunately mistakes happen and we sometimes find ourselves in a situation when we missed an alternative correct answer or a question could be interpreted differently. In those cases the answers need to be adjusted manually (please see limitations for a longer discussion on that). If this is the case, students get the original mark and then a modified mark with an explanation of why an adjustment has been made.

Feedback arrangements
As the marking is done automatically, students receive a pdf file with all of their answers and feedback once the questions and the final score have been verified. The provision of feedback varies. For some questions a simple cue about whether the answer is correct or incorrect is displayed. In some cases the students get an indication of a range – saying that the answer would have been correct if it fell within this range. For some questions, those that are more complex or problematic, a line or two of explanation is provided alongside the indication of an answer being correct or incorrect.

The results of the midterm exam and the mid-term feedback are sometimes discussed in problem classes. Problem classes in this context, as the name suggests focus on students' ability to solve maths problems and this means usually going through

The College encourages staff to create opportunities where students' AI literacy can be developed. College's approach to the use of AI for assessment is outlined in "AI tools in teaching and assessment" document . If the use of AI is prohibited in exam like conditions then students need to be explicitly instructed that is the case and warned about the consequences. the handouts in the class. This creates further

opportunities for learning through clarifying and dialogue. This dialogue takes place with the peers, GTAs and session lead. Due to the nature of those classes it is common to solve mid-term exams there.

The results and solutions are also sent to personal tutors so that they can discuss progress with their tutees. The role of a personal tutor in EEE is twofold, there is a strong pastoral component but there are also opportunities for academic support, i.e. there is scope for solving problem sheets and checking progress across the subject. For example, a student may be provided with problems/ questions from several modules to solve during a personal tutor session. The idea is that if a student is struggling to solve those problems that are relatively very simple compared to the ones they are doing in problem classes and lectures, that's a red flag for the personal tutor that the student might be behind and need more support from the Senior Tutor.

This kind of academic support is also the case with mid-terms. Personal tutors are allocated a week which is dedicated to student progress and the mid-terms. Here the focus shifts away from the technical aspects covered in problem classes towards general discussion about the progress and what is affecting it. The majority of students score well, around 60%-70% but there is a small number of those who score low which can be an indicator of future performance and issues with passing the module.

The data collected by the software (Figure 4) also allows for a higher level feedback to the tutor about the overall cohort performance. The data provide information as to: which question attracted most incorrect answers, average cohort marks, overall marks distribution, highest and lowest scores, easiest and hardest item, participant with highest score etc. This is an advantage of the electronic format and the software as this level of information was difficult to obtain from paper exams.

The process can be very resource heavy and implementation can take a lot of time to get right. Even having done it for the last 4 years we are still in the pilot stage ironing out additional difficulties. Successful implementation across the department requires one person, a digital lead, overseeing the process. This includes, liaising with staff about appropriate question types, inputting the questions and managing potential appeals and adjustments to automatic marking. It is also important to convey the message that while the marking will be easier, preparation of a good mid-term exam will require a lot of effort initially. This effort, however, pays off in the long run.

It is important to note that the time frame in which an electronic exam and a paper exam needs to be prepared are different. Any type of electronic assessments requires more time for quality assurance up front to fix platform related issues and human error, which is much more amplified in the digital space.

Online adaptations

During Covid the pre-term exams were done remotely. In this case collusion was an issue and the degree to which it could be stopped depended on the module leader's engagement in the whole process. If only one version of the exam with a low number of questions (for example 7 which is on the lower side) was provided, this encouraged communication between the students. Collusion is also difficult to prove in a remote setting. You can control for that by for example, having different versions of the same exam (some modules would have 5 sets of different questions), randomizing the order in which the questions appear, observing students' online behaviour and identifying those who do not appear to be answering questions (WISEflow allows that). As a result of that students had the perception of the remote exam not being fair. Perceptions of fairness are much higher when everyone is in the same room this is why post pandemic the mid-terms returned to the in person form but still delivered electronically.

Feedback load in the case of mid term exams is rather small but AI can be used effectively to refine feedback. Feedback notes can be entered into chat GPT and those can be written up as student facing feedback by the software. Specific prompts can be entered around the tone of the delivery of critical feedback.
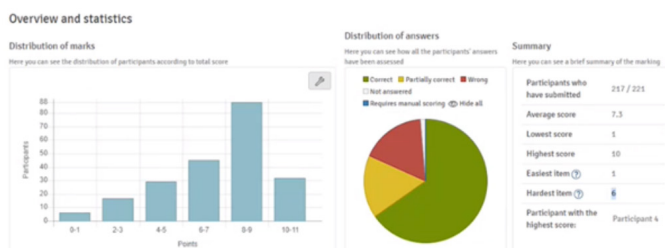

Figure 4 - Exampple of exam data overview

Electronic Mid-term Exams

Advantages of the assessment type

- A mid-term exam is a good way to monitor student progress and offer support for the higher stakes assessment.
- The electronic format allows for an easier way to test large classrooms and provide timely feedback that can actually inform student learning without overburdening the team – the benefits therefore are to both students in terms of quick feedback and the focus on learning and staff in terms of lower marking burden
- The analytics behind the software give an interesting and important insight into student behaviours during the mid-term exam and overall cohort performance. This data can be used to adjust future assessments and identify potential issues.

Limitations of the assessment type

- One of the issues with midterm exams is potentially assessing students when they have not learnt enough. For some courses the first chapters that the first half of the term focuses on are simply introductory and the important analytical skills only come towards the end of the term. This forces the tutors to test some of the knowledge that normally wouldn't be tested in the past as the focus would be on higher order thinking that those initial chapters only lie a foundation for. This makes it easier for the students to score marks for basics that would never be tested in isolation in the past on the end of term exams and makes it easier to pass the course based on the initial knowledge. It can also sometimes create an illusion of a student doing well and passing all of the mid-terms only to fail the higher stake exams.
- When planning a low stakes mid term assessment it is very important to consider if that is an appropriate stage to assess the students at and whether this is the right time to receive formative feedback. If it's not the right time, then it's useful to consider the need for the assessment overall and if there is a genuine need for it then rethink the timing of it.
- When it comes to embedding employability skills, outside of developing knowledge that will be potentially used in future employment

exams of this sort do not naturally lean towards transferable skills. It is important therefore to think where else, potentially around revision such skills could be developed.

- There are some limitations around software and question design/ marking. If we take maths or physics problem that needs to be calculated sequentially, for example, I may first need to calculate the speed of a falling object before using this to calculate its velocity. The results of the later stage depends on the result of previous stages. In automated marking, if a student makes an arithmetical mistake in an initial stage then everything is marked as wrong as the solutions are wrong, perhaps despite correct understanding and application of the process. In a tutor marked examination this can be recognised and if a student makes an early calculation error, but clearly understand the thinking and process they may only lose 1 or 2 marks, whereas in automatic marking they will lose all marks. One suggestion to rectify this is having more intermediate steps to a question. This, however, only works if a mistake is made lower down the problem line. Overall, in numerical questions a small mistake therefore becomes a big mistake.
- The software does not allow for any changes once the exam has been conducted. What can happen is, once the test has been conducted, we find another acceptable form of an answer or discover mistakes in a question would warrant a different answer. Ideally, it would be great to amend an answer so that the mark would be changed automatically for all students. This is unfortunately currently impossible, in such circumstances scores need to be adjusted manually, which is time and resource intensive.

Advice for implementation

- Put in reiterative feedback processes on the questions to eliminate as many mistakes as possible. On a written test when marking you can address potential mistakes with the answers immediately by adjusting the points accordingly but such mistakes are very much visible on an online test and have higher resource repercussion if discovered once the mid-term has been administered (please see the limitations)
- There should be one central person responsible

for managing departmental electronic mid-terms. This helps with consistency and helps to address problems faster. Preferably this person should also be a tutor heavily involved in writing the questions and overseeing the marking because it's really important to understand traditional question types in order to make informed decisions about when to use those in an electronic format. This is an important insight that a learning technologist cannot offer. It is also important to ensure that there are other staff members that are broadly familiar with the systems to ensure continuity in case of illness.

- While we as teachers understand the difference between high and low stake assessments, students sometimes tend to approach both types with the same level of preparation and potentially same level of stress. It is important to spend some time positioning the mid terms as an opportunity to learn and help them adjust their learning and time management accordingly. Watch this video on the importance of achieving good balance and how to manage students emotions.

- Ensure that whatever software you choose is supported by the College. This should be discussed and agreed with your Faculty Learning Technology teams. You don't want to find yourself in a position when you put a lot of manhours into one piece of software to not be able to use it the next year because the College withdraws the funding.

- 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.

- Ensure that the technology you choose to deliver electronic exams allows to make inclusive adjustments to the layout of individual questions and exams

- Ensure that mechanisms are put in place in terms of allowing extra time so that students with learning difficulties are not disadvantaged

- Consider the place of this assessment in the broader context of module and programme design and the ILOs. This will help to identify any points of potential over assessment.

- When choosing the right software it is advisable to liaise with the Ed Tech team as early as possible and use their experience to optimise the tool and the questions and provide instructional support about how to embed the tool for the students throughout the module;

- Quality assurance of electronic exam questions should work on the basis of the more eyes, the better hence a lot of parties should be involved in testing the questions and the answers;

- If fairness is ever an issue, with certain types of software randomisation of of answers is possible (I.e. giving students the same question but with options in different orders for MCQ types of questions or randomize numerical variables within a question).

- Brief students clearly about the rules around the use of AI for any assessment.