CASE STUDY

Authenticity in Learning, Teaching and Assessment

Brendan Masterson, Department of Design Engineering and Mathematics, Middlesex University, London, UK. Email: b.masterson@mdx.ac.uk
Nicholas Sharples, Department of Design Engineering and Mathematics, Middlesex University, London, UK. Email: n.sharples@mdx.ac.uk
Matthew M. Jones, School of Management, University College London, UK. Email: matthew.m.jones@ucl.ac.uk
Snezana Lawrence, Department of Design Engineering and Mathematics, Middlesex University, London, UK. Email: s.lawrence@mdx.ac.uk
Alison Megeney, Department of Design Engineering and Mathematics, Middlesex University, London, UK. Email: a.megeney@mdx.ac.uk

Abstract

Mathematics graduates traditionally are recognised for their problem solving, critical thinking, and analytic skills. However, the methods often used to hone these skills at university are often abstract and decontextualized. This can often create a disconnect between expected capabilities of a mathematics graduate from employers and the actual problem-solving skills required in their career. In this case-study we will outline how the programme team has developed our approach to authentic learning and authentic problems to bridge this gap and ensure graduates are prepared for the workforce.

Keywords: authentic assessment, mathematics education.

1. Introduction

Historically, graduates of degrees in mathematics have garnered recognition for being in considerable demand among employers, attributable to the competencies cultivated throughout the course of their academic studies. The QAA Subject Benchmark Statement (QAA, 2023) summarises these graduate outcomes, describing a typical mathematics degree as enabling students to "*develop graduate attributes which include an enhancement of many general skills, such as numeracy, IT skills, critical understanding and assessment of complex problems, and the ability to identify and analyse problems leading to formulation of solutions, as well as subject-specific skills such as mathematical modelling, data analysis and numerical methods*." However, it is often the case that skills such as critical thinking and problem solving are developed in a more abstract setting than can be communicated to non-mathematicians. Anecdotal evidence, from employers with whom Middlesex University consulted when designing the degree programmes and from students that have graduated from the degree programmes, suggest that what mathematics graduates often lack is evidence of using these problem-solving skills in a real-world setting and an ability to communicate their results to a non-mathematical audience. One particularly useful approach to enable this is the use of authentic assessments. This has been noted in the literature; Pitt and Quinlan (2022) state that "*[a]uthentic assessments can … generate products that showcase students' capabilities more richly than grades alone, making them potentially useful for rethinking how student performance is communicated to [employers]*".

Middlesex University has always emphasised applied, practical skills with a view to ensuring that graduates are prepared for work. The university's stated signature pedagogy is primarily concerned with active, practice-based learning and inclusivity through technology. As a result, the mathematics team have incorporated these philosophies into our BSc Mathematics and BSc Mathematics and Data Science programmes.

Gulikers et al. (2004) describes authentic assessment as "*an assessment requiring students to use the same competencies, or combinations of knowledge, skills, and attitudes that they need to apply in the criterion situation in professional life*".

This is especially difficult to pin down given the diversity of graduate destinations that a typical graduate from mathematics degrees end up in. This, therefore, raises the question: what does authentic assessment mean in mathematics?

In this case study we will present part of an approach taken to answer this question, namely to allow flexibility in the form students can submit their coursework, whether this takes the form of traditional written mathematical coursework, a written discussion, or a multimedia presentation. This approach allows students to enhance their creativity and communication skills while ensuring they build a portfolio of evidence of their skills for potential employers. This is what we will call authentic assessment since the assessment can take the same format as tasks they would be expected to perform in the workplace.

Middlesex University students are loaned iPads and Apple Pencils throughout their degrees, with common software installed on each one. This common platform enables the team to utilise industry standard tools, such as RStudio Server, into our teaching and assessment. Embedding industry tools and the use of real-world data provides students with the opportunity to enhance their problem-solving skills in a more applied setting. Problems fitting this description we will call authentic problems since solving the tasks themselves will involve a similar skillset to those they are likely to encounter in graduate positions.

In this case study we will discuss the approaches we employed to teach and assess mathematics undergraduates in an authentic way. We will make a distinction between authentic assessment and authentic problems in sections 2 and 3 and discuss how this was made possible through the common platform. Section 4 will explore potential approaches to scaling these approaches to larger modules and the difficulties brought on by Large Language Models (LLMs), such as ChatGPT.

2. Authentic assessment

The landscape of assessment in Higher Education has changed drastically since the CoViD-19 pandemic. In response to the necessary move off campus and the inability to have in person examinations many universities began to embrace authentic assessment. Middlesex University in particular, mandated the removal of all end-of-year examinations that were not required by professional bodies for accreditation university wide. The maths team made the decision to remove exams and replace them with authentic assessment at the point the undergraduate mathematics programmes were being revalidated in 2020, prior to the university making the same decision. Authentic Assessment is often defined in a similar way to the Gulikers et al. (2004) cited above. Villarroel et al. (2018) distilled more than 100 articles on authentic assessment published between 1988 and 2015, concluding that authenticity in assessment can be summarised by the following three criteria:

- **Realism**. Authentic assessment emphasises the practical relevance of problems to the professional setting, communicating assessment briefs more closely aligned to the language graduates might encounter in the workplace.

- **Challenging**. Assessments might be less well-formed, requiring students to demonstrate higher-order thinking skills such as problem solving, critical thinking and critical understanding.
- **Self-evaluative**. Transparency, for example, of marking criteria allow students to develop self-reflection and self-evaluation.

Middlesex University internal guidance similarly defines authentic assessment as having the following characteristics:

- Is realistic;
- Requires judgement and innovation;
- Asks students to ‘do’ the topic;
- Replicates or simulates the context in which adults are tested in the workplace;
- Assesses the student’s ability to efficiently and effectively use a repertoire of knowledge and skills to negotiate a complex task;
- Allows opportunities to rehearse, practise, consult resources, and get feedback.

When we began to incorporate authentic assessment into our programme we had to do so in a way that respected the university’s definition of authentic assessment. Our interpretation of the above definition was that students should be given the opportunity to enhance their employability skills while still testing their mathematical ability. What resulted was pieces of coursework in which students could choose the format in which they could submit their coursework. Many of these formats reflect the kind of task graduates may have to complete in the workplace, including:

- Traditional written coursework;
- Written discussions;
- Multimedia submissions.

Staff may encourage students to consider submitting in one of these formats. This approach had a number of advantages. First it gave students the opportunity to decide if they want to target an employability skill as part of a piece of coursework, and if so, which of these they would like to work on. Second it promotes inclusivity among students by allowing them to select the format of submission that they believe best reflects their mathematical ability. To help aid consistency of marking across formats, and to remain true to the philosophy that assessment should reflect workplace tasks, there will be marks available for the overall presentation of submissions in any format. Marks for the presentation of a submission regardless of format prioritises the clarity, correctness and mathematical rigour of the problem-solving method rather than good video editing skills or the use of graphics. We do however appreciate knowledge of appropriate software for submissions, for example LaTeX for written submission. This has the added benefit that this reduces the need to make reasonable adjustments for the assessment since it is built in. Finally, it allows the students to exercise their creativity in selecting their format of submission.

These innovations were welcomed by students, with one student even using one of their multimedia as part of a successful job application which was cited by their employer as a highlight of their application. More detail of this work can be found in the paper Masterson et al. (2022).

3. Authentic Problems

Authentic problems are distinct to authentic assessment, or at least our implementation of authentic assessment. An authentic problem is a problem which is similar in format and context to the kind of task a student can expect in the workplace.

Section B

Work with your lecturer to identify a topic that is of interest to you, and find a dataset related to this topic that is appropriate to analyse.

Question 1: Provenance (Total 2 marks)

Describe your dataset.

- How trustworthy is it?
- How accurate is it?
- What could be done to improve the quality of the data?

Question 2: Data analysis (Total 18 marks)

Write R code to analyse the dataset, calculate appropriate statistics, and interpret your findings.

In particular you should demonstrate your knowledge and understanding of *summary statistics* and *confidence intervals*.

Your code should be commented to explain what it does in mathematical and statistical terms. You should refer to the mathematical and statistical theory of the course to justify any claims you make or assumptions about the data.

Show a draft to your lecturer by *Friday 17th November*.

Your lecturer will then set a follow-up question that should be addressed in your final submission (see Question 3).

Question 3: Personalised follow-up question (Total 10 marks)

This question will be set by your lecturer after reading your draft work. It will involve combining data from an additional dataset.

Question 4: Large Language Model comparison (Total 5 marks)

Use a Large Language Model (such as ChatGPT) to attempt to write a data analysis similar to your solution to Question 2. (We will discuss possible prompts in the workshops). Copy and paste the prompt and output of the Large Language Model.

How does your data analysis compare? In what ways is it more useful?

Figure 1. Example of authentic assessment from a mathematical statistics course

It is important to note that an authentic problem is not simply an applied problem. An applied problem may be very well defined and therefore it is readily apparent to the student what techniques they are expected to utilise to solve the problem.

An authentic problem on the other hand will often:

- be vaguely defined;
- require experimentation;
- allow students to use every tool at their disposal;
- obligates students to attend structured guidance sessions.

This accurately reflects what will often be expected of mathematics graduates in the workplace, particularly those whose work colleagues do not have strong backgrounds. For example, one can imagine that it would not be unusual for a maths graduate working as a data scientist to be asked to analyse a dataset by their project leader or line manager with no further instruction. In this scenario, they would likely do some descriptive analysis and probe the data before reporting their findings to their manager seeking further guidance on what analysis the manager would like them to perform. These situations are likely very common for many of our graduates. Therefore, our assessment should be preparing them for situations just as these.

On our programmes there is a dedicated Level 5 Problem-Solving and Communication module where students gain familiarity with the nature of these open-ended problems in class. Thus, students have the confidence to tackle these open-ended problems when they begin to encounter them in summative assessment later in the academic year.

Figure 1 shows an example of an authentic problem from our second year mathematical statistics module.

This assessment closely reflects what a student may be expected to do if they're asked to analyse a dataset in the workplace. They must make some explicit judgement on the quality and source of the data. The brief is open ended. They must send draft to their lecturer to receive guidance and a follow-up question, replicating the employee manager dynamic. Finally, they must compare their report to that produced by a LLM such as ChatGPT. This serves two purposes: first it will discourage the student from using an LLM to produce their own report. More importantly, this is also authentic since a job applicant may be expected to demonstrate why their work is superior to that produced by AI.

While statistics may readily lend itself to this approach, it is still possible in pure mathematics modules. Figure 2 shows another example from a multivariable calculus module.

In figure 2 we see many of the same features again: judgement required, open-ended brief, and opportunities for consultation. Furthermore, they are again instructed to make a comparison to what is produced by ChatGPT. If a student were to choose a banana for this assignment and compare their solution to that produced by ChatGPT they would obtain output similar to that shown in figure 2.

This parameterises the surface in figure 4. Clearly this is not a banana! Students witnessing output such as this should be more aware of the limitations of LLMs and hesitant about using is to wholly complete their own coursework.

4. Larger Modules

There is the question of how such method would scale to larger modules. The method of having lecturers as project managers would seem to be quite demanding on staff time at a time when staff have increasingly large workloads. One of the advantages of in person exams, despite their inauthenticity, is that they are incredibly time-efficient for staff.

The received wisdom we got from an experienced colleague was that examinations are 10 times faster to marks than to sit. If we were to take this ratio as a given the below assessment breakdown would represent the same workload of 18 minutes per student for staff:

- 180 minute written exam;
- 120 minute exam and 6 minute viva;
- 60 minute exam and 12 minute viva;
- Three vivas of 6 minutes.

Allowing for more vivas would present more opportunities for authentic assessment reflecting the workplace with the lecturer playing the role of project manager. Vivas have long been used to assess students understanding of material and naturally lend themselves to this approach. Furthermore, past research from the team has shown that our students generally prefer having conversations with their lecturers to assess their ability over end of year examinations (Masterson et al., 2022).

It is this approach that we plan to utilise on one of our service modules in the forthcoming academic year. The team administer a service module for the Business School in Middlesex. The module typically has 120 students enrolled on it. Previously this module was assessed with a two-hour exam with sections A and B. Assuming our colleague's ratio these should take 12 minutes to mark each. Generally, the questions were closed, quite prescriptive and did not allow for a lot of problem solving.

Considering the university removing non-professional body examinations from all programmes the assessment for this module now needs to be revised. What we are proposing is a hybrid approach. Our assessment will still have two sections. Section A will have closed questions and prescriptive and therefore test similar skills to the previous incarnation of the assessment. Section B will be a 6-minute viva which will test the student's problem-solving skills. This should be a format that would not additionally to staff workload while incorporating the principles of authentic assessment even for large modules.

Real world calculus

We will look at an applied problem using multi-variable calculus.

1. Look around campus or at home for an object to model in 3D. The object should be curved and have no sharp edges so that we can model it using a continuously differentiable surface parametrisation.

   Now, use the "GeoGebra 3D Calculator" App to define a surface parametrisation for the surface of your chosen object.

   > **Pro-tip**
   >
   > Finding a good parametrisation is an iterative process - you'll have to try lots of things!
   > Try to think mathematically about your object:
   >
   > - Is the object 'built from' a curve?
   > - Is your object a collection of circles?
   >
   > The commands in https://wiki.geogebra.org/en/Surface_Command will be really useful.

   Write a log of your thinking as you develop your parametrisation. This log should describe how you are building up the parametrisation from your observations of the object.

   Bring some draft work to the workshop in Week 10. Your lecturer can support you with this activity.

   Your submission should include:

   - A photo of your chosen object (using your iPad).
   - A .ggb file that contains the parametrisation.
   - The surface parametrisation written down using appropriate mathematical notation.
   - The log of your thinking as you developed the parametrisation.
   - A short video of your chosen object with the parametrisation overlayed on top. This can be done using the "Augmented Reality" (AR) feature of GeoGebra and making a screen recording of the iPad.

   15 marks

2. Use an optimisation technique to find the greatest value of the z-coordinate for your parametrisation. 5 marks

3. What is the surface area of your parametrisation? Write down an appropriate integral that would give the surface area, and simplify as much as possible.

   It will almost certainly be impossible to evaluate this integral by hand. You should use numerical methods (e.g. from https://www.wolframalpha.com/) to produce the final answer. Include the code that you use to produce the numerical solution. 5 marks

4. The university plans to build an MRI scanner nearby your object. This means that your object will be subjected to strong magnetic field **B** described by

$$\mathbf{B}(x, y, z) = \begin{pmatrix} 0 \\ y \\ -z \end{pmatrix}$$

   Write down an appropriate integral to calculate the total magnetic flux through your object, and simplify as much as possible.

   It will almost certainly be impossible to evaluate this integral by hand. You should use numerical methods (e.g. from https://www.wolframalpha.com/) to produce the final answer. Include the code that you use to produce the numerical solution. 5 marks

Figure 2. Example of authentic assessment from a calculus course

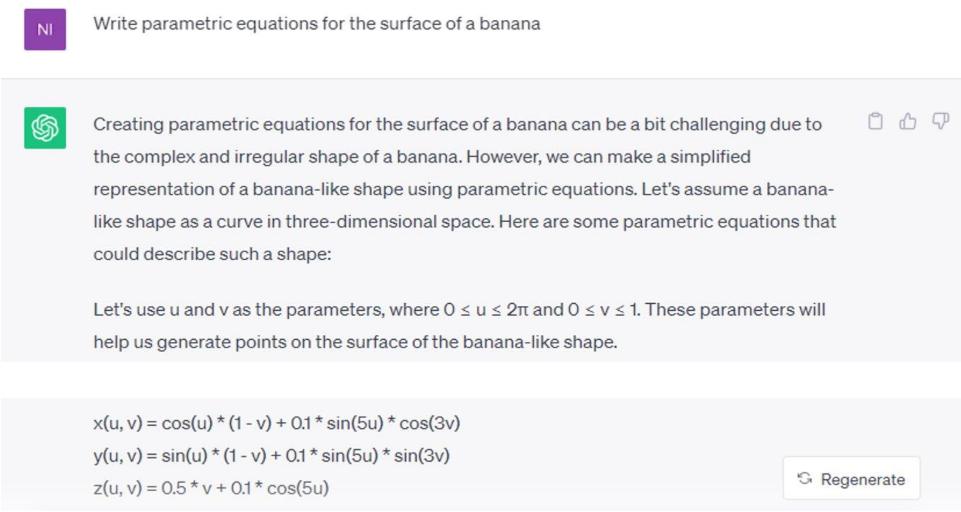

Figure 3. Extract from the dialogue of a ChatGPT request

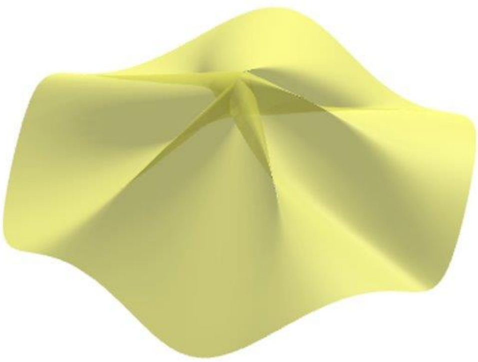

Figure 4. A ChatGPT generated banana

5. Conclusion

This case study has outlined the maths team's approach to authentic assessment and authentic problems to support a student's problem-solving skills and more adequately prepare them for the workplace.

Building flexibility into assessment encourages students to think creatively about how to best approach the problem while negating the need for reasonable adjustments. Importantly this flexibility uses formats which are typical of a professional workplace, motivating students to target formats which will enhance their employability skills. This approach is made possible because we know that each student has access to identical hardware and software.

Authentic problems reflect the kind of problems they will encounter in the workplace. Typically, these problems are vaguely defined and open-ended. These problems will require students to meet and consult with their lecturer who will act as a project manager. Lecturers will play an active role as the students complete their assessment by guiding students on expected content to satisfy module learning objectives and issuing bespoke follow-up questions making the assessment dynamic.

The ready availability of LLMs presents a great difficulty for universities. Our view is that rather than discouraging student from using LLMs we should, where possible, encourage students to actively compare their results with the output of an LLM. This will make students more aware of the limitations of LLMs, thus allowing them to articulate why their work is superior to that produced by an LLM. This is an authentic real-world problem since these students may have to justify why they should get a position instead of the potential employer simply using an LLM.

While this form of authentic assessment and problems could be extremely beneficial in preparing students for the workplace it is important to balance how these techniques scale against staff workload. However, with careful planning we believe that it is possible to implement this approach at scale with adding significantly, if at all, to the workload of staff. This is subject to future work of the maths team.

6. References

Gulikers, J.T.M, Bastiaens, T.J. and Kirschner, P.A. (2004). A Five Dimensional Framework for Authentic Assessment. *Educational Technology Research and Development*, 52(3), pp.67-86. https://doi.org/10.1007/BF02504676

Masterson, B., Jones, M., Megeney, A., Sharples, N., and Lawrence, S. (2022). Authentic by Design: Developing Mathematicians for the Talent Economy. *MSOR Connections*, 21(2), pp.44-51. https://doi.org/10.21100/msor.v21i2.1400

Pitt, E. and Quinlan, K.M. (2022). *Impacts of higher education assessment and feedback policy and practice on students: a review of the literature 2016-2021*. York: Advance HE. Available at: https://www.advance-he.ac.uk/knowledge-hub/impacts-higher-education-assessment-and-feedback-policy-and-practice-students-review [Accessed 3 April 2024].

QAA [The Quality Assurance Agency for Higher Education] (2023). *Subject Benchmark Statement: Mathematics, Statistics and Operational Research*. Available at: https://www.qaa.ac.uk/the-quality-code/subject-benchmark-statements/subject-benchmark-statement-mathematics-statistics-and-operational-research [Accessed 13 November 2023].

Villarroel, V, Bloxham, S, Bruna, D, Bruna, C and Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, 43(5), pp.840-854. https://doi.org/10.1080/02602938.2017.1412396