MSOR Connections

Articles, case studies and opinion pieces relating to innovative learning, teaching, assessment and support in Mathematics, Statistics and Operational Research in higher education.

Volume 22 No. 3.

Special Issue: CETL-MSOR Conference 2023 Part 2

Contents

For information about how to submit an article, notifications of new issues and further information relating to *MSOR Connections*, please visit https://journals.gre.ac.uk/index.php/msor.

This journal is published with the support of the **sigma** network and the Greenwich Maths Centre.

EDITORIAL

Peter Rowlett, Department of Computing, Sheffield Hallam University, Sheffield, UK.
Email: p.rowlett@shu.ac.uk

Welcome to this issue of *MSOR Connections*, the second to collect papers from the CETL-MSOR Conference 2023 held at Cardiff University on 7th-8th September 2023. I mentioned the longstanding connection and shared history between that conference and this journal in the previous issue's editorial, noting that both emerged from the Maths, Stats and OR Network (in the case of the conference, in collaboration with the various Centres for Excellence in Teaching and Learning associated with MSOR subjects). The MSOR Network was a Subject Centre of the Learning and Teaching Support Network/Higher Education Academy from 2000-2012. After the HEA closed its subject centres, *MSOR Connections* was relaunched by the HEA itself in 2013, but only lasted two issues in that format (Vol. 13 No. 1 and 2).

In 2014, a group of us got together and agreed a plan to revive *MSOR Connections* as a community initiative, with support from the **sigma** Network and University of Greenwich. This group formed the first editorial board and the first set of editors (Noel-Ann Bradshaw, Alun Owen, Rob Wilson, and me). We agreed an editorial workflow that would allow us to publish the journal on a sustainable basis with no budget, with at first four (more recently three) editors taking turns to be responsible for an issue – organising peer review of articles, making editorial decisions, compiling and copyediting the issue, and writing the editorial. We published our first issue (Vol. 14 No. 1) in autumn 2015 and the current issue (Vol. 22 No. 3) is the twenty-sixth since the revival, representing around one third of the total issues of *MSOR Connections* published in its quarter century so far.

This is the tenth issue of *MSOR Connections* for which I have been responsible since the relaunch[1], and it will be my last. I have had the privilege of publishing issues with 81 articles featuring the work of 153 authors from ten countries[2] (17 authors appear in more than one of these issues; one author appears in four of them!). I have twice worked with guest editors, chairs of conferences or workshops from which the issue collects papers. I've also enjoyed creating the images for the covers since volume 19, taking over from Rob Wilson who did it until then. It has all been a delight, and I am pleased to have played a small role in the history of this journal. I leave it in the safe hands of my fellow editors Tony Mann and Alun Owen.

But first, this issue opens with an interesting opinion article exploring the use of language in statistics and whether this is a barrier to understanding for students – and those supporting them! – who are not fluent in statistical English. Rollings opens a conversation about what we can do to help. Another issue in supporting students in using statistics is test selection. Goodale tries out ChatGPT as an interactive alternative to flowcharts for students to get self-help on the right test to use.

Moving onto teaching practice, Kuveke, Shaker and Prendergast report on the design and delivery of a multidisciplinary first-year undergraduate statistics module, taking into account the common interests and disciplinary differences of the student cohorts. Shearman, Aziz and Pettigrew describe the development and analysis of an online diagnostic tool as part of a university mathematics refresher program. Using a hierarchical knowledge map, this tool attempts to reduce the number of questions students must answer to determine their level of understanding. Brewer draws inspiration from a career in pre-university education in an attempt to make large-group teaching at university more interactive. Fairfax and Saha attempt to use a group project to improve the sense of cross-cultural student community on a degree where a large number of overseas students join in year 2.

Connections between mathematics support and other academic support programmes are explored by Hurley and Casey, whose new delivery model involves closer collaboration which has highlighted possible links between support connected to mathematics anxiety and mathematical resilience and support offered through coaching.

Finally, Ardito explores the post-pandemic transition from a blended learning course to a traditional face-to-face format, sharing an 'anthology' approach to repurposing digital materials as a supplementary resource.

MSOR Connections is a venue for our community to share its innovative practice, and also a place for those new to writing and reviewing articles to get started. We have some experience supporting new authors and new peer reviewers. Though I am leaving the editorial team, I intend to continue writing and reviewing articles, knowing only too well that the journal could not operate without people like you and I doing both. I encourage you to join me – submitting articles or signing up with your reviewing interests via the journal website https://journals.gre.ac.uk/index.php/msor.

1. 14(2), 15(2), 16(3), 18(1), 18(2), 20(1), 20(2), 21(3), 22(2) and 22(3).

2. UK and Ireland, of course, plus Australia, Finland, Germany, Italy, Japan, Netherlands, South Africa, and USA.

OPINION

Talking Statistics: A reflection on some of the problems with statistical language

Lois Rollings, Learning Enhancement Team, Middlesex University, London, UK.
L.Rollings@mdx.ac.uk

Keywords: statistics, language.

1. Introduction

For most of my life I managed to swerve statistics. I learned a little at school and studied some (very theoretical) statistics as part of my mathematics degree. As a teacher, in schools and then university, I did not teach anything beyond GCSE statistics. It was not until I got a post in Mathematics and Statistics support a decade ago that I had to begin to learn the subject properly.

I had excellent support from the **sigma** network, attending a memorable SPSS Bootcamp and other events which helped me enormously. But I was conscious that there were various aspects of statistics that presented problems for me. One was the way in which statistics differed from mathematics in being much less cut and dried. If a student had made a mistake in a calculation or argument it was fairly easy to spot and correct. However, when a student said 'my supervisor said I should do a t-test' and this did not seem the most appropriate way forward it was much harder to advise. I also realised that I was finding the language of statistics harder to master than I felt it ought to have been. It is this aspect that I will focus on in this article as I tentatively suggest that students might also have such problems.

2. Statistical Language

Language acquisition is not a trivial process (Leung 2005, cited by Kaplan et al., 2010, p.1) and specialized language can lead students to perceive the subject as more difficult than it is (Lemke 1990, cited by Kaplan et al., 2010 p.2). The language of a discipline can be seen as "*a tool for participation in a community of practice*" (Espinet et al., 2012, cited by Kaplan and Rogness, 2018, p.1) and it is surely part of our job to induct students into that community.

2.1 Statistical English and Lexical Ambiguity

Rangecroft (2002) discusses how Mathematical English (ME) differs from Ordinary English (OE) and suggests that Statistical English (SE) is distinct from both. I considered myself fairly fluent in OE and ME but it was with SE that I had difficulties.

One issue is that statistics often uses words from OE but gives them a particular technical meaning. This can give rise to 'lexical ambiguity' (Barwell 2005, cited by Kaplan 2011, p.56). An everyday example of this is that if you are reading a sentence starts 'The bat was…' and this brings to mind a cricket or baseball bat but then the sentence continues 'roosting in the belfry', it will take a moment for you to change your perspective. It has been shown that this disambiguation requires higher levels of brain activation than reading a sentence where there is no ambiguity (Mason and Just, 2007, p.115).

There are many terms used in SE that have a meaning in OE but where the statistical meaning differs from the everyday meaning. In some cases the two meanings are close; in others the difference can be quite large. Some examples are 'significant', 'random' and 'spread' and you can doubtless think of others. This means that students might start to think of the everyday meaning and might take a while to readjust to the statistical meaning. Perhaps a more serious concern is that they might not realise that the technical meaning is different and so persist with the everyday meaning. There is also the added complication that words used in both mathematics and statistics, for instance ‘normal’ and ‘estimate’, may not have the same meaning in both contexts.

Konold 1995 (cited by Kaplan et al., 2010, p.2) reports that statistics students have “*strongly-held, but incorrect, intuitions that are highly resistant to change*”. This brings to mind the Piagetian ideas of ‘assimilation’ and ‘accommodation’; it is much easier to assimilate an idea into an existing schema than to have to reconfigure the schema (Mason and Johnston-Wilder, 2004, pp.149-151). The problem is of course compounded if as teachers we do not realise that what we are talking about is not what is being understood by the students, and this issue may occur in other subjects. This means that “*there is a vast amount of room for misunderstanding that may never be detected*” (Roth, 2005, cited by Kaplan and Rogness, 2018, p.1).

The term ‘random’ is a good example of this. Perhaps particularly with younger students the everyday meaning of random is ‘weird’. Kaplan et al. (2014, p.12) found that “*students persist in thinking that random processes are haphazard, weird or have unlikely outcomes*” and describe using two images in their teaching to try to address this. A picture of some people dressed as zebras illustrated the everyday meaning and a hat illustrated the technical statistical meaning. They report some success in using this strategy as throughout the statistics course they could remind students that they meant the “*hat not the zebras”* (Kaplan et al., 2014, p.13).

In some case it is the formation of a ‘nominal group’ that provides a technical meaning, as in ‘standard deviation’ and ‘degrees of freedom’. Some of these groups are ‘dense’, for instance ‘Pearson’s product-moment correlation coefficient r’. Different selections of these words can be found in different sources and exactly what is meant can differ. From a linguistics perspective the question is which of these is the ‘head noun’. Are we talking about the abstract idea of correlation or the coefficient? As McConway (2016) remarks “*product has nothing to do with the production of goods and moment has nothing to do with time*”.

The issue of lexical ambiguity is not confined to mathematics and statistics. Rector et al., (2013). discuss its presence in scientific disciplines and evolutionary biology in particular. It is also not peculiar to English; Ilana Lavy & Michal Mashiach-Eizenberg (2009) describe similar issues in Hebrew. As well as there being an overlap between OE and SE, there may also be an overlap with technical terms used in other disciplines. In accounting ‘variance’ just means difference and I have been confused by this myself. ‘Sample’ might have a different meaning for biomedical students (Dunn et al., 2016, p.14). It might seem counterintuitive but lexical ambiguity might be less of a problem for students whose first language is not English as they may not be so familiar with the everyday meanings of words (Dunn et al., 2016, p.10).

Another potential problem is that statistical language is not standardised (Dunn et al., 2016, pp.9-10). Variables might be described as ‘predictor’, ‘explanatory’ or ‘independent’. Different software packages also use different terminology; Minitab refers to a ‘2-sample’ t-test whereas SPSS uses ‘independent sample’ t-test. This can be a particular issue for those of us working in a support context as we will not always know which term the student is used to. The same word can also be used with different meanings. For instance, ‘average’ might refer to the mean (as it does in Excel) or as a term for one of several averages (mean, median or mode). An average might also be referred to as a ‘measure of centre’ or a ‘measure of central tendency’ and students might not realise that these are all the same thing. Even in the name of the discipline itself there is a distinction between the study of ‘statistics’ and a ‘statistic’.

2.2 Statistician's Names and Definitions

Something I found particularly confusing at first was the profusion of statistician's names attached to different tests or calculations. For some reason it took me a long time to distinguish the Mann-Whitney and Kruskal-Wallis tests. (And embarrassingly much longer to discover that in each case the names refer to two people not one with a double-barrelled surname.) In some SPSS output there will be several options (all named) for a given test and differing opinions on which one should be used. In some ways I should be used to this as many mathematics theorems are 'named'. But although the same theorem might have different names attached to it perhaps out of national pride (see the inequality named after, variously, Cauchy, Schwarz and Bunyakovskii) there will not be alternative versions of the same theorem. This highlights one of the key differences between mathematics and statistics; once a mathematical theorem is accepted as true, it will not be challenged, whereas in statistics there might be no standard way of doing a calculation.

One thing which still strikes me as odd is that it seems very rare for sources to explain who these people were. Of course, there are other issues, beyond the scope of this article, in that several statisticians of the early twentieth century were eugenicists. This has been acknowledged by some institutions, for instance University College London has 'denamed' buildings (UCL, 2020), but not, as far as I know, in statistics textbooks.

Although many statisticians' names appear, statistical writing can seem impersonal. Use of the passive "*obscures the human agency involved"* (Morgan 2005, p.110) and increases 'social distance'. (This was a linguistic term long before it was used in relation to a pandemic.) I suggest that this might be a particular problem for students from social sciences. As Graves et al. (2014, p.7) remark *"[definition] performs different roles among disciplines"*. In social sciences and lower level maths a definition might formalise a pre-existing concept. An author might cite several examples and then say 'We call this X'. In mathematics, certainly at university level, it is usual to start with a definition, 'X is defined as…' and this formulation is often used in statistics. In statistics definitions might not even always be agreed (Dunn et al., 2016, p.17).

3. Concluding Remarks

A key question is what we can do about these issues but, just as there are different ways of calculating, say, an effect size, there is no consensus on the approach that we should adopt.

Rangecroft (2002, p.36) encourages "*trying to pre-empt difficulties by careful choice of language*" and suggests that we should "*become more attuned to the possibilities of misunderstandings arising from language difficulties*". Kaplan and Rogness (2018, p.4) concur, saying that one of the first steps is for "*instructors to recognise words that have lexical ambiguity*".

Some argue for a more consistent use of technical terms. Kaplan et al. (2011, pp.59-60) advocate using the term 'variability' rather than 'spread' but warn not to expect "*a change in terminology to be a panacea*". Kaplan et al. (2010 p.17) suggest using 'measures of centre' instead of 'averages' but, as Dunn et al. (2016, p.20) remark, it would be "al*most impossible*" to get consistency across sources and we need to respect different usage in different disciplines.

While Lavy and Mashiach-Eizenberg (2009, p.7) argue that technical terms should be introduced using both formal and informal language, others, for instance Neibert et al. (2012, cited by Kaplan and Rogness, 2018, p.3), say that using common everyday words as an alternative to technical terms might increase misunderstandings and barriers for students. Dunn et al. (2016, pp.20-24) discuss at some length the different solutions offered by others and conclude that a balance needs to be struck "*between the mathematical and linguistic requirements"* of a statistics course, noting that how this is achieved will depend on the context.

Finally, it is perhaps worth noting that our language use, particularly when speaking, is not conscious (Martin, 2010, p.14). So, even if we could agree on an approach, we will not always get this right.

4. References

Dunn, P.K., Carey, M.D., Richardson, A.M. and McDonald, C. (2016). Learning the language of statistics: Challenges and Teaching Approaches. *Statistics Education Research Journal,* 15(1), pp.8-27. Available at: http://iase-web.org/Publications.php?p=SERJ [Accessed 7 November 2023].

Graves, H., Moghaddasi, S. and Hashim, A. (2014). "Let *G=(V, E)* be a graph": Turning the abstract into the tangible in introductions in mathematics research articles. *English for Specific Purposes,* 36, pp.1-11. https://doi.org/10.1016/j.esp.2014.03.004

Kaplan, J., Rogness, N.T. and Fisher, D.G. (2010). Lexical Ambiguity in Statistics: How Students Use and Define the words: Association, Average, Confidence, Random and Spread. *Journal of Statistics Education,* 18(2), pp.1-17. https://doi.org/10.1080/10691898.2010.11889491

Kaplan, J., Rogness, N.T. and Fisher, D.G. (2011). Lexical Ambiguity: making the case against spread. *Teaching Statistics,* 34(2), pp.56-60. https://doi.org/10.1111/j.1467-9639.2011.00477.x

Kaplan, J., Rogness, N.T. and Fisher, D.G. (2014). Exploiting lexical ambiguity to help students understand the meaning of *random. Statistics Education Research Journal,* 13(1), pp.9-24. Available at: http://iase-web.org/Publications.php?p=SERJ [Accessed 7 November 2023].

Kaplan, J. and Rogness, N.T. (2018). Increasing Statistical Literacy by Exploiting Lexical Ambiguity of Technical Terms. *Numeracy,* 11(1), Article 3. https://doi.org/10.5038/1936-4660.11.1.3

Lavy, L. and Mashiach-Eizenberg, M. (2009). The Interplay Between Spoken Language and Informal Definitions of Statistical Concepts. *Journal of Statistics Education,* 17(1). https://doi.org/10.1080/10691898.2009.11889502

Martin, J. (2010). Language, register and genre. In C. Coffin, T. Lillis and K. O'Halloran, eds. *Applied Linguistics Methods: a Reader*. London: Routledge in association with the Open University, pp.12-32.

Mason, J. and Johnston-Wilder, S. (2004). *Fundamental Constructs in Mathematics Education,* Abingdon: RoutledgeFarmer.

Mason, R. and Just, M.A. (2007). Lexical ambiguity in sentence comprehension. *Brain Research,* 1146, pp.115-127. https://doi.org/10.1016/j.brainres.2007.02.076

McConway, K. (2016). More confusing terms in statistics. *Open Learn*. Available at: https://www.open.edu/openlearn/course/view.php?id=12784 [Accessed 7 November 2023].

Morgan, C. (2005). Words, Definitions and Concepts in Discourses of Mathematics, Teaching and Learning. *Language and Education,* 19(2), pp.103-117. Available at: https://discovery.ucl.ac.uk/id/eprint/1564242/ [Accessed 7 November 2023].

Rangecroft, M. (2002). The Language of Statistics. *Teaching Statistics,* 24(2), pp.34-37. https://doi.org/10.1111/1467-9639.00080

Rector, M.A., Nehm, R.H. and Pearl, D. (2013). Learning the Language of Evolution: Lexical Ambiguity and Word Meaning in Student Explanations. *Research in Science Education*, 43, pp.1107-1133. https://doi.org/10.1007/s11165-012-9296-z

UCL (2020). *UCL denames buildings named after eugenicists*. Available at: https://www.ucl.ac.uk/news/2020/jun/ucl-denames-buildings-named-after-eugenicists [Accessed 1 November 2023].

RESEARCH ARTICLE

Using Generative AI to help with statistical test selection and analysis

Tom Goodale, Academic Achievement Team, Liverpool John Moores University, Liverpool, UK.
Email: t.r.goodale@ljmu.ac.uk

Abstract

One of the most common questions that students ask statistics advisors is 'What test should I do?' This paper explores the use of generative AI chatbots, specifically ChatGPT, as a tool to assist students, in particular those with limited experience in statistics, in selecting appropriate statistical tests for their analyses. Traditional methods, such as flowcharts and online test selectors, require at least a basic understanding of measurement scales and research design, which can be an issue for many students who have limited exposure to statistics on their courses. This research focuses on developing and refining prompts to guide ChatGPT in providing accurate and relevant statistical test recommendations. A hypothetical scenario was used to test the effectiveness of various prompts, ranging from simple, naïve questions to more sophisticated ones utilising specific prompt patterns, such as the 'context manager' and 'flipped interaction.' These patterns were selected to enhance the chatbot's responses and ensure the relevance and accuracy of the test suggestions. The findings suggest that while AI chatbots like ChatGPT can be a valuable resource for students, their effectiveness is highly dependent on the quality of the prompts used. The paper concludes with a discussion on the potential of these AI tools in educational settings, acknowledging the limitations of current technology and suggesting directions for future research and development.

Keywords: statistical test selection, ChatGPT, generative AI, prompt engineering.

1. Introduction

Choosing the correct statistical test is a core task of any statistical analysis. Whilst experienced users of statistics will not often struggle with this task, those with less experience, especially students, often struggle with this. In order to help with this many textbook authors, such as Field (2018), include some sort of flowchart to help in the decision making process; a simple web search finds a plethora of flowcharts available online. Additionally, students may be pointed towards one of the free online statistical test selectors (Jackson-Wood, 2023; Sage, 2023; Dokuz Eylül University, 2023; Sheffield Hallam University, 2023), one of which was developed after much research (Suner, Karakülah and Dicle, 2014), or they may be pointed towards commercial ones such as on the Laerd statistics website (Laerd, 2023).

In order to use these flowcharts and online test selectors successfully, students need to know something about measurements scales, i.e. the distinction between nominal, ordinal and interval/ratio (Stevens, 1946). Additionally, many of the flowcharts and automated test selectors start with questions around the specific research hypothesis that is being tested – commonly a distinction is made between a test of association, a test of group differences, or prediction – although some flowcharts and test selectors instead start by asking what the measurement levels are. Thus, when teaching students how to use these tools it is important to teach both the measurements scales and, at least at a coarse level, something about study design. Many students do not have this necessary background and thus have to ask supervisors or statistical support staff to help them choose a test.

The increasing sophistication of generative AI chatbots such as ChatGPT potentially provides a powerful alternative to this traditional method of choosing a statistical test, and could, to a large extent, reduce the need for students to seek external help.

At the current time chatbots do, however, require some care to use, especially for naïve users. The quality of the response is highly dependent on the quality of the prompt, and even with the best prompts there is still the possibility of 'hallucinations' where the chatbot confidently provides information which is factually incorrect. This paper explores prompts and reports an effective prompt that naïve statistical users might use for test selection.

2. Chatbot Prompts

Prompts are at the heart of the interaction between a user and an AI chatbot. The simplest prompts ask a question which the chatbot answers by use of its pre-trained material. More sophisticated prompts go beyond using a single question and answer; they can introduce extra material to supplement the chatbot's own data, establish a tone and style for the chatbot's responses, set a context within which a question is set, or control the format of output. Recently White et al. (2023) have taken the idea of 'design patterns' from software engineering (Gamma, 1995) to categorise common 'prompt patterns.' These patterns, detailed in Table 1, offer a structured approach to interacting with chatbots such as ChatGPT.

Table 1. Prompt patterns from White et al. (2023).

Pattern Category	**Prompt Pattern**
Input Semantics	Meta Language Creation
Output Customization	Output Automater Persona Visualization Generator Recipe Template
Error Identification	Fact Check List Reflection
Prompt Improvement	Question Refinement Alternative Approaches Cognitive Verifier RefusalBreaker
Interaction	Flipped Interaction Game Play Infinite Generation
Context Control	Context Manager

Of particular interest to this paper are the 'context manager' pattern, which allows context to be set, and the 'flipped interaction' pattern, which asks the chatbot to query information from the user, and the 'persona' pattern, which helps to shape the interaction.

3. Developing a prompt for Statistical Test Selection

In order to develop prompts and test their effectiveness, the following hypothetical scenario was used:

- An experiment was conducted to test the effectiveness of five different fertiliser concentrations on crop height.
- This was performed via a 10x5 design whereby ten fields were each split into five plots. In each field the five plots were each assigned one of the five levels of fertiliser.
- 20 measurements of crop height were taken in each plot.

In this experiment the response variable is the crop height and there are two explanatory factors – the amount of fertiliser, which is the factor of interest, and the field, which should be considered a blocking factor. This experimental design should lead to the recommendation of a two-way ANOVA.

All tests were done with ChatGPT 3.5, which is freely available. The results may have been different with the more powerful ChatGPT 4.0, or with other AI chatbots. As I was running these tests I tried to answer as I would expect a student to do so from my experience but may not always have succeeded in this.

3.1 Naïve prompt

As a baseline test a naïve prompt of ‘What test do I need to do to examine the relationship between crop height and fertiliser’ was used. This is the sort of question I might expect to see from a student coming to me.

The response from ChatGPT was (see Supporting data, Example 1):

> To examine the relationship between crop height and fertilizer, you can conduct a statistical analysis using various methods, depending on the nature of your data and your research goals. Here are some common tests and approaches you can consider:
>
> **Correlation Analysis**:
>
> **Pearson's Correlation Coefficient**: Use this if both crop height and fertilizer are continuous variables, and you want to assess the strength and direction of the linear relationship between them.
>
> **Spearman's Rank Correlation**: Use this if your data is not normally distributed or if there might be non-linear relationships between the variables.
>
> **Regression Analysis**:
>
> **Simple Linear Regression**: This is appropriate if you want to predict crop height based on the amount of fertilizer used. It provides a linear equation that describes the relationship between the two variables.
>
> **Multiple Regression**: If you have more than one independent variable (e.g., multiple types of fertilizers or other factors influencing crop height), multiple regression can be used to analyze their combined effects.
>
> **Analysis of Variance (ANOVA)**:
>
> **One-Way ANOVA**: Use this when you have multiple groups of crops that received different levels or types of fertilizer, and you want to determine if there are statistically significant differences in crop height among the groups.

Two-Way ANOVA: This is used when you have two independent variables (e.g., type of fertilizer and time of application) and you want to assess their individual and interactive effects on crop height.

Chi-Square Test:

If your data is categorical and you want to examine the relationship between discrete levels of fertilizer and crop height categories (e.g., tall, medium, short), you can use a chi-square test for independence.

Non-parametric Tests:

If your data violates assumptions of normality or homoscedasticity, consider non-parametric tests like the Mann-Whitney U test or Kruskal-Wallis test.

Time Series Analysis:

If you are collecting data on crop height and fertilizer application over time, you might use time series analysis techniques to examine trends, seasonality, and autocorrelation in the data.

Data Visualization:

Before conducting formal statistical tests, it's often a good idea to visualize your data using scatter plots, box plots, or other relevant plots to gain insights into the relationship between crop height and fertilizer.

The choice of test depends on your specific research design, the type of data you have, and the research question you are trying to answer. It's essential to carefully plan your study, collect data systematically, and consult with a statistician or data analyst if you are unsure about the appropriate analysis method for your specific situation.

This is a very broad answer; it has introduced statistical concepts that the student may be unfamiliar with, and whilst the correct test is there, it has not singled it out. A student receiving such an answer might be overwhelmed and could potentially spend a lot of time learning about irrelevant analyses before narrowing it down to ANOVA and subsequently two-way ANOVA.

3.2 First attempt at a more sophisticated prompt

In order to overcome the shortcomings of the naïve prompt, three prompt patterns suggested themselves:

- To shape the interaction and make it more relevant the persona pattern was introduced via 'Act as a statistics adviser'
- To reduce the complexity and assumed background the context manager pattern was introduced via 'talking to an undergraduate student' and 'The student wants to know what statistical tests to do with their data.'
- The flipped interaction pattern was then used to force ChatGPT to ask questions of the user in order determine the test: 'Ask questions one at a time until you can recommend a test and then provide your recommendation.'

Thus, the full prompt was

Act as a statistics advisor talking to an undergraduate student with no knowledge of statistics. The student wants to know what statistical tests to do with their data. Ask questions one at a time until you can recommend a test and then provide your recommendation.

For reasons of brevity the full ChatGPT output is not reproduced here, but can be accessed online (see Supporting data, Example 2).

With this prompt ChatGPT asked a sequence of questions and provided the following recommendation (see Supporting data, Example 2):

> Great, if your crop height measurements are normally distributed within each group and have roughly equal variances, you can proceed with a one-way ANOVA test. This test will help you determine if there are statistically significant differences in crop heights between the different levels of fertilization.
>
> In your case, you have five different levels of fertilization (including the control group with no fertilizer). You can set up the null hypothesis (H0) as: “There is no significant difference in crop heights between the different levels of fertilization.”
>
> And the alternative hypothesis (H1) as: “There is a significant difference in crop heights between at least two of the fertilization levels.”
>
> Running a one-way ANOVA test will provide you with an F-statistic and a p-value. If the p-value is less than your chosen significance level (typically 0.05), you can reject the null hypothesis and conclude that there is a statistically significant difference in crop heights between at least two fertilization levels.
>
> If the ANOVA test indicates significance, you can perform post-hoc tests (e.g., Tukey's HSD) to identify which specific fertilization levels are different from each other.
>
> Remember to report the results, including the F-statistic, p-value, and any post-hoc test findings, in your research report or paper.

Whilst not the ideal test, as it doesn’t consider the blocking effect of the fields, it is a concrete test and ChatGPT has also gone some way to suggest ways in which it could be reported.

3.3 An improved prompt

Whilst the previous prompt was a marked improvement on the naïve prompt, ChatGPTs occasionally asked several questions at once, which could lead to confusion and might not always be fully answered by students. In order to try to counteract this, the prompt was modified to:

> Act as a statistics advisor talking to an undergraduate student with no knowledge of statistics.
>
> I want to know what statistical tests to do with my data.
>
> Ask questions one at a time until you can recommend a test and then provide your recommendation. The questions must be short and explain any statistical terms used.
>
> Ask your first question.

This succeeded in its goal of forcing ChatGPT to ask shorter, more direct and specific questions (see Supporting data, Examples 4, 4c, 4d, 4e). In many cases, but by no means all, this also led to the recommendation of a 2-way ANOVA.

3.4 Tests with other scenarios

This prompt was also used for two other scenarios:

- Spring extension vs weight: ChatGPT suggested regression, which would be appropriate; with the information given in the scenario, however, correlation would also have been appropriate and simpler (see Supporting data, Example 5).
- The relationship between gender (nominal with 2 categories) and car colour (nominal with 2 categories). ChatGPT correctly identified a Chi-squared test of independence for this case, and

was also able to answer questions on assumption checking when prompted (see Supporting data, Example 6).

4. Conclusions

AI chatbots such as ChatGPT provide a powerful alternative to more traditional methods of statistical test selection such as flowcharts and interactive test selectors. With effective prompts they can be used by statistically naïve users to make sensible test selections, thus allowing these users to make a start on their analysis without having to speak to supervisors or statistics advisors. These recommendations, however, may not be the ideal tests, and may not use the full structure available in the data, and indeed may change from run to run. Ideally students should still have a certain level of statistical knowledge in order for them to judge the quality of the answer. One possible strategy they may choose to adopt is to run the prompt several times and then investigate the different tests that are recommended in order to better choose the best one; generative AI may also be of use for this step as it would be able to compare and contrast the tests for the given scenario.

Free generative AI chatbots such as ChatGPT 3.5 have the potential to give students swift access to statistical help at times when it may be difficult for them to access a statistics advisor, such as overnight, at weekends, or at particularly busy times of year. Whilst this advice is not perfect, they would allow students to progress their analysis and complete their assignments. As the sophistication of chatbots increases the quality of advice will almost certainly increase.

In the future the community could develop more sophisticated prompts for the free versions of chatbots. The recent release of 'Custom GPTs' by Open AI may also provide a way for the community to build a custom statistics advisor GPT which could be made available via a web application.

5. Supporting data

Example 1: https://chat.openai.com/share/28b6b6bb-b5d9-4cf3-8d6d-6b90e0a86dfa

Example 2: https://chat.openai.com/share/5e4b7330-b09a-4fc1-b3c2-caf0cc25cfa7

Example 4: https://chat.openai.com/share/8d95b63e-d9c1-4452-be04-ff3d1e74723d

Example 4c: https://chat.openai.com/share/cdb4b9be-8259-4e9c-91c8-6f2e43e2a6ed

Example 4d: https://chat.openai.com/share/f94af8e8-77e6-4e83-8034-30092d134904

Example 4e: https://chat.openai.com/share/61c23440-41b5-4a6c-ab9a-f76d139e88bc

Example 5: https://chat.openai.com/share/341a61bc-371c-4b0a-a018-42afb6e2d0f5

Example 6: https://chat.openai.com/share/204004fe-4cf0-4aad-aa7c-fcad99fc7525

6. References

Dokuz Eylül University (2023). *StatXFinder: a decision support tool for appropriate statistical test selection*. Available at: https://medinfo.deu.edu.tr/statxfinder/index.php [Accessed 13 November 2023].

Field, A.P. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Los Angeles: SAGE.

Gamma, E. (1995). *Design patterns: elements of reusable object-oriented software*. Addison-Wesley.

Jackson-Wood, M. (2023). *statistical test flowchart*. Available at: https://www.statsflowchart.co.uk/ [Accessed 13 November 2023].

Laerd (2023). *Laerd Statistics*. Available at: https://statistics.laerd.com/ [Accessed 13 November 2023].

Sage (2023). *Which Stats Test*. Available at: https://methods.sagepub.com/which-stats-test [Accessed 13 November 2023].

Sheffield Hallam University (2023). *Interactive test chooser*. Available at: https://maths.shu.ac.uk/mathshelp/WhichTestIndex.php [Accessed: 20 December 2023].

Stevens, S.S. (1946). On the theory of scales of measurement. *Science*, 103 (2684), 677-680. https://doi.org/10.1126/science.103.2684.677

Suner, A., Karakülah, G. and Dicle, O. (2014). Towards a web-based decision support tool for selecting appropriate statistical test in medical and biological sciences. *Studies in Health Technology and Informatics*, 205, pp.48-52.

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J. and Schmidt, D.C. (2023). A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT. *arXiv preprint*. https://doi.org/10.48550/arXiv.2302.11382

CASE STUDY

Designing an introductory statistics subject for students with diverse educational backgrounds and chosen qualifications

Rupert E. H. Kuveke, Department of Mathematical and Physical Sciences, La Trobe University, Melbourne, Australia. Email: r.kuveke@latrobe.edu.au
Amanda J. Shaker, Department of Mathematical and Physical Sciences, La Trobe University, Melbourne, Australia. Email: a.shaker@latrobe.edu.au
Luke Prendergast, Department of Mathematical and Physical Sciences, La Trobe University, Melbourne, Australia. Email: luke.prendergast@latrobe.edu.au

Abstract

This is a case study on the design of a first-year undergraduate statistics subject at La Trobe University, entitled Making Sense of Data, which is taken by students from various disciplines. To account for students' diverse educational backgrounds and chosen qualifications, this subject is designed such that all students complete core statistics concepts, while a third of the subject contains stream-specific content. This subject design provides students with a solid foundation in statistics, while addressing the demand for a flexible first-year statistics subject which is accessible and relevant for students enrolled in a variety of tertiary degrees. This structure allows for stream-specific lectures, computer lab material, assessments, and even statistical software programs to be used across different streams. The design also incorporates strategies for addressing statistics anxiety within the curriculum. In this paper, we present the outcomes of this subject design in terms of student performance, engagement and satisfaction. We also present iterative and reflective changes that have been made to the subject over time, in response to student and staff feedback, and discuss the impact these changes have had on student outcomes.

Keywords: higher education, statistics anxiety, flexible curriculum, student engagement, undergraduate statistics.

1. Introduction

Statistical literacy is an increasingly vital and necessary skill in today's data-driven world. In recognition of this, a wide variety of higher education courses often contain a compulsory introductory statistics component in their curriculum (see e.g. Tishkovskaya and Lancaster, 2012). While the demand for quality tertiary-level statistics teaching continues to increase, numerous challenges in the teaching and learning of statistical content remain (Bromage et al., 2022). Students undertaking a first-year undergraduate statistics subject (known as a course or module in UK terminology) may have diverse educational backgrounds and chosen tertiary qualifications, and indeed for many students, such a subject may be their only formal tertiary-level exposure to statistics.

As a result, these students can often fail to appreciate the importance or relevance of statistics to their chosen degree(s), and can approach learning statistics with low motivation and negative attitudes (see e.g. Gal and Ginsburg, 1994; Conners, McCown and Roskos-Ewoldsen, 1998). Indeed, learning statistics has been argued to be similar to learning a second language (Lalonde and Gardner, 1993), with all the attendant difficulties. This can be compounded by the fact that large percentages of students can also experience statistics anxiety (Onwuegbuzie and Wilson, 2003), a long-acknowledged form of anxiety (see Cruise, Cash and Bolton, 1985) which is related to, but distinct from, mathematics anxiety (see e.g. Zeidner, 1991, Paechter et al., 2017). Statistics anxiety can lead to a lack of enjoyment of and engagement with the content, reduced knowledge retention, and lower pass rates (see e.g. Macher et al., 2012, González et al., 2016 and Marshall et al., 2022). These poor attitudes towards learning statistics may in part be promoted by prior negative experiences, with Koparan (2015, p.103) noting that there are *“serious problems experienced in teaching and learning statistics”* in schools.

Given these challenges, in order for a first-year undergraduate statistics subject to be successful, it must be flexible enough to account for students’ various educational backgrounds, and be able to address and mitigate barriers to learning such as statistics anxiety, while also ensuring students receive a solid foundation in statistical knowledge, regardless of their chosen tertiary qualification(s).

In recognition of these requirements, a new first-year La Trobe University subject entitled *Making Sense of Data* (with subject code STM1001) was designed during 2021 and first offered in 2022. STM1001 consists of a core statistics module (about two-thirds of the subject), and stream-specific modules (about one-third of the subject) which focus on the use of statistics in specific disciplines. Students complete both the core module, and a stream-specific module which best fits their chosen qualification. This structure ensures that students receive a solid background in foundational statistics concepts, while also learning statistical concepts which more closely align with their chosen qualifications. This novel design structure facilitates content delivery to large cohorts of students from numerous disciplines, while being flexible enough to allow for the learning materials, assessments, and statistical software programs used to differ between streams. Similar approaches that cater for diverse cohorts do exist: for example, by using examples relevant to all disciplines and/or discipline-specific tutorial support (e.g. Hilliam and Calvert, 2019, Calvert et al., 2022). As well as including these elements in its design, STM1001 additionally includes stream-specific lectures, assessment and use of statistical software packages in its design.

In this paper, we present a case study on the creation, design and implementation of STM1001, and explain the pedagogical rationales behind our design processes, decision-making and delivery styles. We then present an evaluation of the subject’s success, in terms of student performance and satisfaction, based on its operation since teaching began in 2022. The design and implementation of STM1001 is an ongoing, dynamic process, and so we also detail some recent modifications made to the subject, in response to student and staff feedback, along with preliminary evaluation of these changes. We conclude by discussing the current state of STM1001, and our future intentions.

2. Background

2.1 Motivation from an institutional perspective

In early 2020, the University sought to reduce duplication of teaching efforts into subjects with overlapping content that were offered either independently within different areas (e.g., different departments and disciplines), or within the same area with content tailored to the needs of students from within different degrees. The key objectives were to (i) create greater efficiency whereby workloads associated with duplication could be reduced, (ii) implement targeted efforts for continuous subject improvements in a single subject to ensure that each student’s educational needs are met and (iii) ensure that subject content was created and delivered by discipline experts.

Led by the Pro Vice-Chancellor Learning & Teaching in the then called College of Science, Health and Engineering, the discipline of statistics was identified as part of this effort. At that time, the Department of Mathematics and Statistics (now the Department of Mathematical and Physical Sciences), offered four introductory statistics subjects: Statistics for Psychology, Statistics for the Life Sciences, Statistical Methods, and Statistical Science. Further, introductory statistics was also taught, either in whole, or in part, in many other departments, e.g., in the health sciences and in business. An initial stakeholder meeting was held in May, 2020, and was attended by the two Pro Vice-Chancellors of Learning & Teaching from the university’s two colleges, the Head of Department of Mathematics and Statistics, and academic representatives from science and business disciplines. The meeting agreed on two things. Firstly, there was substantial overlap between various introductory statistics subjects across the university. Secondly, that there were discipline-specific needs within many of those subjects.

Concurrently during 2020, the university was undergoing extensive course re-restructuring in alignment with a guiding set of principles called *Course Architecture*. These principles provided structural parameters for all courses that provided greater opportunities for subjects, minors, majors and specialisations to be shared across courses. Course Architecture also included the provision for a subject to be constructed out of two or more modules. For example, a typical single subject is 15 credit points (8 subjects make up a full-time load for one year). According to a set of rules and where it is a suitable to do so, a 15 credit point subject could be made up of modules which collectively contribute to the student's learning needs.

Following the initial meeting, and in full consideration of Course Architecture, the Department of Mathematics and Statistics proposed a university-wide modularised introductory statistics subject consisting of a common core module for all, as well as stream-specific modules to cater for the different needs of students across different degrees and qualifications. This proposed modularised structure meant that new modules could be created and included in the future to suit new needs (e.g., new qualifications or identified gaps in existing qualifications). This proposal was agreed to by stakeholder disciplines and development started in 2021.

2.2 Core and Stream-specific design

STM1001 content has been designed to ensure that (a) all students develop a strong understanding of foundational statistical concepts which are broadly applicable, while (b) also learning statistical skills which relate more directly to their chosen area(s) of study. As explained in the previous section, this is achieved by the use of modules. All students complete the core module, which accounts for 70% of the subject content and assessment. Students also complete one stream-specific module, which accounts for the remaining 30% of the subject content and assessment. The robust architecture of the streams allows for significant differences between streams, such as the use of different software and assessments, while still being flexible enough to support crossover of content where relevant.

Currently, two stream-specific modules, the Science/Health (SH) module and the Data Science (DS) module, are offered, with the potential for further modules to be added in future. The SH module is for students enrolled in Health Sciences-, Life Sciences-, Applied Life Sciences- and Social Sciences-focused courses, while the DS module is for students enrolled in Statistics-, Mathematics- or Information and Communications (ICT)-focused courses. Initially, the primary software package used for all students was the R software environment for statistical computing and graphics, via the RStudio integrated development environment, in accordance with stakeholder requests. Further discussion on statistical software packages used is provided in section 2.5.

Stream-specific content and software skills are assessed via two stream-specific assignments that students complete throughout the teaching period. These stream-specific assignments contain unique questions relevant to the specific stream, but utilise the same questions and data sets across streams where possible (i.e. where content overlaps between streams, see figure 1), for efficiency. Each stream-specific assignment contributes 15% to a student's overall mark for STM1001. The remaining 70% of assessments are stream-agnostic and are completed by all students.

Figure 1 outlines the different STM1001 core and stream-specific topics covered throughout the teaching period. The topics included were chosen following extensive consultation with key stakeholders, including program coordinators of Science and Health degrees whose students would be taking STM1001. Following consultation, it became clear that the needs of students in both Health and Science degrees were very similar, which led to the Science/Health stream being designed as one module rather than two separate modules. The core module covers typical introductory statistics content including descriptive statistics and plots, t-tests (one-sample, paired and independent samples), ANOVAs, linear regression, and chi-squared tests. In the stream-specific modules, SH students focus predominantly on study design, while DS students focus on R coding, data visualisation and machine learning. The specific examples and data sets discussed in each of these streams have been carefully selected, to ensure they are relevant to the students (see e.g. Cornock, 2016). This supports students in developing an appreciation of the benefits of statistics in the context of their chosen tertiary qualification. For stream-specific Topics 7 and 8, students from both streams cover the same material.

Topic	Core (All Students)	Science/Health Stream	Data Science Stream
1	Introduction to Statistics and presenting data	Why Research?	R Basics
2	Descriptive Statistics	Research Questions and Ethics	Data Visualisation I
3	Probability and Distributions	Designing a Study I	Data Visualisation II
4	Sampling Distributions	Designing a Study II	Data Visualisation III
5	Hypothesis Testing, One-sample t-tests	Designing a Study III	Simulations in R
6	t-tests for two-sample hypothesis testing	Designing a Study IV	Writing R Functions
7	One-way ANOVA	Big Data I: Cluster Analysis	
8	Correlation and Simple Linear Regression	Big Data II: p-value Adjustments	
9	Hypothesis Testing for One and Two Sample Proportions	Repeated Measures Analysis	Machine Learning I
10	Chi-squared Tests for Categorical Data	Reading Research	Machine Learning II
11	Statistical Power and Sample Size Calculation	Hierarchies of Evidence	Machine Learning III
12	Revision	Revision	Revision and AI Introduction

Figure 1. An overview of the current STM1001 core and stream-specific content covered throughout a teaching period.

2.3 Statistics Anxiety

Previous research has highlighted that 70% to 90% of students may experience anxiety when confronted with the prospect of learning statistics (Zeidner, 1991; Marshall et al., 2021). When designing STM1001, a concerted effort was made to integrate measures to address and mitigate statistics anxiety. Sessions to discuss and normalise the phenomenon of statistics anxiety are built into the early weeks of the subject curriculum, to help defuse students' concerns, and to highlight that they are not alone in feeling anxiety about learning statistics. Raising awareness of statistics anxiety early in the subject allows students to begin considering methods for tackling this issue, as advised by e.g Marshall et al. (2017). These interventions provide opportunities to encourage and reassure students about their learning (Wilson, 1999), and reinforce the message that anxiety about statistics and mathematics is a valid and common scientific phenomenon (see e.g. Lyons and Beilock, 2012). These interventions also serve as an opportunity to remind students of the support services that are available to them. We provide weekly one-on-one and small group support sessions with experienced tutors, following e.g. Lalayants (2012). In addition, students have access to the Maths Hub (see e.g. Jackson, 2021), a La Trobe University student support initiative which provides free mathematics and statistics support to all La Trobe University students throughout the university's teaching periods.

A number of other elements recommended in the statistics anxiety literature (e.g. Chew and Dillon, 2014) have been incorporated into the design. These include: less emphasis on mathematical calculations and more emphasis on conceptual understanding and use of software; weekly low-stakes quizzes that have generous time limits and multiple attempts available; humour and enjoyable in-class activities including classroom polling (e.g. Kahoot https://kahoot.com/ and Mentimeter https://www.mentimeter.com/); immediacy behaviours in class such as using students' names and being approachable; collecting data in class; and student support provided by personalised emails sent via the Student Relationship Engagement System (SRES, Lui et al., 2017).

STM1001 is taught both in blended and fully online instances. While students in online subjects may experience lower levels of statistics anxiety on average than those learning in face-to-face settings (see e.g. DeVaney, 2010; Marshall, Mahmood et al., 2022), online learning presents its own challenges, and can be an isolating experience. To help reduce statistics anxiety and foster a sense of community and belonging (see Lear, Ansorge and Steckelberg, 2010; O'Shea, Stone and Delahunty, 2015), we host a weekly Zoom session for STM1001 online students. To promote interaction between students and build rapport between staff and students, these sessions are used to review current subject information, discuss content-related issues, and revise recent content in a fun, low-stakes atmosphere using Kahoot quizzes (see Shaker, Hurst and Marshall, 2021; Russell, 2022).

2.4 Open access materials

STM1001 content was designed to maximise interactivity and open access for all students. Learning resources are provided as open access material on RPubs.com, bookdown.org, and YouTube, and can be accessed via links provided on the subject Learning Management System (LMS). Topic readings, lectures, supplementary guides, computer lab content, assignments and the exam were all written in R v4.0.4+, using packages such as bookdown, RMarkdown, xaringan, and exams (Grün and Zeileis, 2009).

There are numerous advantages to providing access to subject materials on these platforms, including the ability to update and re-publish content in real time, and seamless integration of videos into online readings and computer labs, as well as interactive elements (Dunn, Brunton and Farrar, 2022). In addition, the open access availability of materials leads to them having a wider reach, contributing to the public good, as well as the opportunity for readers to provide feedback. For example, since 1 January 2022, the STM1001 reading materials have been accessed by over 50,420 users from 188 countries around the world.

2.5 Iterative changes

We regularly implement iterative changes to STM1001 content and delivery methods, based on teaching experiences and student and staff feedback. In the first semester of teaching, STM1001 used interactive workshops rather than lectures, following a 'flipped learning' approach (Bergmann and Sams, 2014). However, feedback suggested students required more structured sessions when learning new statistical concepts, and so in subsequent teaching periods, lectures on both core and stream-specific content have been gradually added to the subject curriculum. Additional short video guides have also been created to support students in learning key statistical concepts. While these changes have meant an increase in weekly contact hours, feedback on these changes has been positive overall. We have balanced this for students by reducing the amount of content in readings and supplementary material over time, and by replacing some written sections of content with short videos.

We have also made adjustments to statistical software packages used over time and in response to student and staff feedback. Originally, SH students learnt R, in accordance with requests from stakeholders. However, despite additional support resources being created for SH students in the second teaching period, feedback and teaching experiences suggested that for this cohort, jamovi would be a more suitable choice of statistical software package. As a result, in 2023 R was replaced by jamovi in the SH stream, with an immediate improvement in student experience observed. The flexible design of the subject has meant that while SH students use jamovi throughout the subject (core and stream-specific content), the DS students use R throughout the subject, thus optimising the learning experience for all students.

Finally, to support student success, at-risk students are now emailed earlier in the semester so that contact is made and to advise regarding support services they can access. In addition, all students are contacted earlier in the semester advising marks required in remaining assessments to pass the subject. As discussed in section 2.3, all emails are personalised, and sent via SRES (Lui et al., 2017).

2.6 STM1001 Students

The STM1001 cohort is diverse not only in terms of students' educational backgrounds and chosen qualifications, but also location, being taught at three different campuses as well as online. During the four semesters offered so far, 1754 students have completed the subject, as shown in table 1.

Most students (1546) were in the SH stream, while the DS stream has been taken by 208 students so far. Figure 2 shows a further breakdown of students by discipline group over time. Most notably, the number of students from the Health Sciences increased dramatically starting from Semester 1, 2023, at which time STM1001 became a core subject in the Bachelor of Health Sciences.

Table 1. Overview of STM1001 students by stream.

Stream	**Semester 1, 2022**	**Semester 2, 2022**	**Semester 1, 2023**	**Semester 2, 2023**	**Total**
Data Science	19	62	44	83	208
Science/Health	126	392	464	564	1546
Total	145	454	508	647	1754

3. Evaluation

In this section, we provide an evaluation of the subject in terms of the core and stream-specific design, followed by a more general evaluation, particularly with respect to iterative changes that have been made to the subject over time.

3.1 Core and Stream-specific design

A total of 547 comments have been received from institutional Student Feedback on Subject (SFS) surveys carried out over the four teaching periods. 27 of these comments referred directly or indirectly to the core and stream-specific design of the subject. Of these comments, 13 may be considered positive, 13 negative, and one neutral. The main theme arising from the positive comments was that students appreciated learning statistics within a context related to their chosen degree, for example: *"I found the content to be interesting, in how it involves stats and science as a whole. I can see how the skills learnt can be applied later on in my course and or career".* One student also commented on the benefit of studying with peers who are studying towards similar degrees: *"… it was nice to be grouped with students studying similar degrees to me".* The main themes arising from the negative comments were related to not finding the content to be interesting, as well as disadvantages to having the subject split into two parts. For example, *"The two parts of the course for Health students made it feel like doing 2 separate subjects in one".*

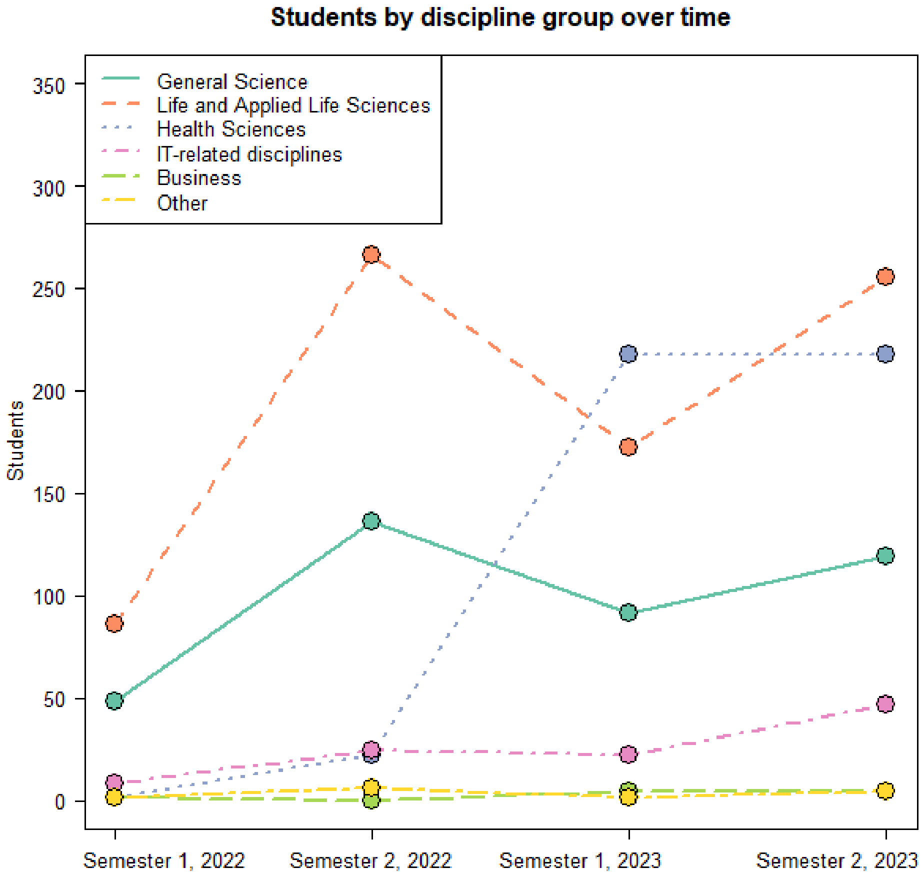


Figure 2. Students enrolled in STM1001 over time by discipline group.

Overall, students appear to benefit from the stream-specific design of the subject, as they are more able to see the relevance of the statistical content they are learning to their chosen degrees. Further benefit may also be gained by further integration of core and stream material, or by making the links between the two more explicit.

As discussed in section 2.1, from an institutional perspective, there are clear efficiency gains, since all introductory statistics can be taught within one subject rather than 2-3 subjects or even up to 6 or more. Some of the efficiency gains made by the institution are offset by the increased academic workload required to run a subject that contains multiple streams. Part of this academic workload increase includes the administration load of placing students into streams and stream-specific classes. In addition multiple sets of curriculum and assessment must be maintained for approximately half of the subject, due to the stream-specific modules and the stream-specific software used throughout the subject in both core and stream classes. Overall though, despite the challenges and additional complexities, the flexible nature of the design affords worthwhile benefits.

3.2 Iterative changes

As discussed in section 2.5, the main iterative changes made to STM1001 over time have included replacing R with jamovi for the SH cohort, gradual incorporation of lectures for less reliance on readings and self-guided learning, reduction of content, and earlier intervention for at-risk students.

Figure 3 shows STM1001 pass rates over time. More specifically, figure 3A shows that the overall pass rate for all students increased from Semester 1, 2022 (64%) to Semester 2, 2022 (71%), and then decreased slightly in Semester 1, 2023 (70%), before increasing again in Semester 2, 2023 (78%). However, as noted previously, STM1001 did not become a core subject in the Bachelor of Health Sciences until Semester 1, 2023, at which time there was a sharp increase in the number of students from this course. Therefore, for a more accurate comparison, the purple line shows the pass rate of the STM1001 cohort excluding students from the Bachelor of Health Sciences. This comparison shows an increase in pass rate over time, starting with 64% in Semester 1, 2022, followed by 71% in Semester 2, 2022, 77% in Semester 1, 2023 and 79% in Semester 2, 2023.

Figure 3B shows the pass rates over time separated by stream, and shows that in general, the pass rate for the DS stream is slightly higher than that of the HS stream.

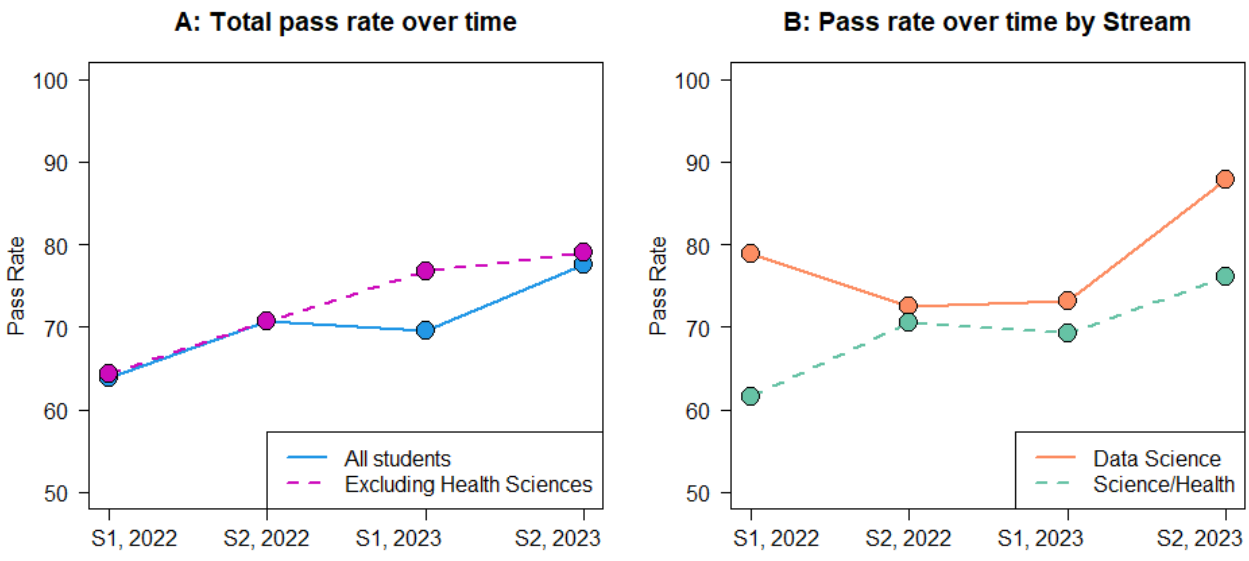


Figure 3. STM100 pass rates over time.

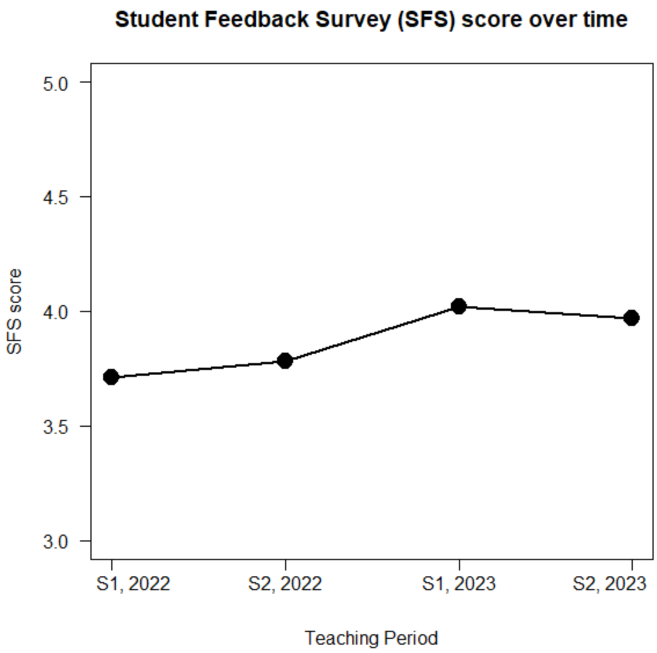


Figure 4. STM100 student feedback over time. Mean score (out of 5) in response to the question, *“Overall I was satisfied with the quality of this subject”.*

In terms of student satisfaction, figure 4 shows a general increase over time. The chart displays data from institutional Student Feedback on Subject (SFS) surveys. In particular, it shows the mean score (out of 5) in response to the question, *"Overall I was satisfied with the quality of this subject".* In Semester 1, 2022, the mean response was 3.71, followed by 3.78 in Semester 2, 2022, 4.02 in Semester 1, 2023 and 3.97 in Semester 2, 2023.

The evaluation of both student performance and student satisfaction over time seems to indicate that the iterative changes made over time have had a positive impact on the student experience. While it is difficult to assign causation to particular changes, anecdotally, the change from R to jamovi for the SH students from Semester 1, 2023 was significant and resulted in an immediate improvement in the student experience. This has been evidenced by a reduction in email traffic regarding technical issues, less reliance on support services for statistical software-related questions, and 16 student comments from the SFS survey specifically mentioning positive experiences with jamovi. A representative student comment from the SFS survey expresses this sentiment: *"I did this subject last year and failed. This year I've really enjoyed aspects such as the transition to jamovi for Sci/health students as I had mainly struggled with the production of graphs previously but this is made for better with the new program"*.

4. Discussion and Conclusions

STM1001 was designed to meet the demand for a robust yet flexible first-year undergraduate statistics subject which could be taught to students with diverse educational backgrounds and chosen tertiary qualifications. The subject covers both typical introductory statistics material, and stream-specific content, in order to provide students with a solid foundational understanding of statistics, while also contextualising the use of statistics in the students' chosen qualifications. The novel design of the subject allows for the use of different statistical software, learning materials and assessments across the streams, while also supporting synergistic learning between the core and stream-specific content, whereby students' learning and adaption of skills in each module is enhanced and contextualised by the learning they have attained in the other.

STM1001 has now been offered for four semesters. Over that time, student performance and satisfaction has generally increased as iterative changes have been made in response to student feedback and teaching experience. Demand for the subject continues to grow, with the STM1001 student cohort in Semester 2 2023 being the largest yet, with a total of 647 active students across the three campuses and online.

We will continue to implement evidence-based, student-centric modifications in order to streamline the subject and to ensure it continues to address the diverse needs of the student cohort and institutional stakeholders.

For future teaching periods, we are considering adjustments including (i) combining the weekly interactive workshops and lectures into an integrated, interactive-style lecture, (ii) increasing the amount of short instructional videos for asynchronous learning, and (iii) merging the weekly 1-hour core and 1-hour stream-specific computer labs into a single combined 2-hour class. We are also planning on modifying assessments to promote greater weighting of early assessments, and to ensure they continue to provide a range of authentic learning experiences. These adjustments are intended to support students' learning, performance and satisfaction and to help mitigate statistics anxiety, and are based on student and staff feedback, recent literature (see e.g. Cox & Taylor, 2018, Thomas et al., 2019, McArthur, 2023) and in consideration of efficiency gains from an academic workload perspective.

Designing, implementing and delivering a tertiary education subject like STM1001 involves a range of challenges. These include developing one subject with a design that is sufficiently complex and flexible such that it can meet multiple disciplinary and student needs, as well as staffing and workload challenges. As such, we offer the following reflections which may prove useful to others embarking on developing a similarly structured subject.

From an institutional perspective, clear and effective communication and discussion with stakeholders is vital to ensure the process and end product align with stakeholder expectations and the overarching strategic vision of the university. Staffing requirements must be considered, to ensure academic workloads are feasible and reasonable for subject and instance coordinators who face the added load required to run a subject that contains multiple streams. Training, support and guidance for large teaching teams within the subject is also key to the subject running successfully. Members of the teaching team may be predisposed to teaching in one specific stream, and this must be taken into account before each period of teaching to ensure staffing numbers are adequate. Assessments must be designed to effectively meet the subject intended learning outcomes and be relevant and engaging to students from a variety of disciplines, while also being scalable and robust to academic integrity violations. Fundamentally, careful planning of the overall subject design, creation and presentation of content, and day-to-day teaching that acknowledges and accounts for complex requirements such as those discussed here, can lead to a subject that successfully caters to students with diverse educational backgrounds and chosen tertiary qualifications.

5. References

Bergmann, J. and Sams, A. (2014). *Flipped Learning Gateway to Student Engagement*. Eugene: ISTE.

Bromage, A., Pierce, S., Reader, T. and Compton, L. (2022). Teaching statistics to non-specialists: challenges and strategies for success. *Journal of Further and Higher Education*, 46(1), pp.46-61. https://doi.org/10.1080/0309877X.2021.1879744

Calvert, C., Hilliam, R. and Steele, E. (2022). Can the same statistics module be used for service teaching by tailoring the support based on the student's chosen qualification? *MSOR Connections*, 20(3), pp.73-83. https://doi.org/10.21100/msor.v20i3.1344

Chew, P.K.H. and Dillon, D.B. (2014). Statistics anxiety update: Refining the construct and recommendations for a new research agenda. *Perspectives on Psychological Science,* 9(2), pp.196-208. https://doi.org/10.1177/1745691613518077

Conners, F.A., McCown, S.M. and Roskos-Ewoldsen, B. (1998). Unique Challenges in Teaching Undergraduate Statistics. *Teaching of Psychology*, 25(1), pp.40-42. https://doi.org/10.1207/s15328023top2501_12

Cornock, C. (2016). Teaching mathematics to Business and Enterprise students in a module based around Excel. *MSOR Connections*, 15(1), p.22-27. https://doi.org/10.21100/msor.v15i1.315

Cruise, R.J., Cash, R.W. and Bolton, D.L. (1985). Development and validation of an instrument to measure statistical anxiety. *Proceedings of the 1985 Statistical Education Section of the American Statistical Association, Las Vegas, NV, USA*, pp.92-98.

DeVaney, T.A. (2010). Anxiety and Attitude of Graduate Students in On-Campus vs. Online Statistics Courses. *Journal of Statistics Education*, 18(1). https://doi.org/10.1080/10691898.2010.11889472

Dunn, P.K., Brunton, E.A., and Farrar, M.B. (2022). Your online textbook is ready: a shareable, interactive online textbook in response to COVID-19 lockdowns. *International Journal of*

Mathematical Education in Science and Technology, 53(3), pp.582-593. https://doi.org/10.1080/0020739X.2021.1983051

Gal, I. and Ginsburg, L. (1994). The Role of Beliefs and Attitudes in Learning Statistics: Towards an Assessment Framework. *Journal of Statistics Education*, 2(2). https://doi.org/10.1080/10691898.1994.11910471

González, A., Rodríguez, Y., Faílde, J.M. and Carrera, M.V. (2016). Anxiety in the statistics class: Structural relations with self-concept, intrinsic value, and engagement in two samples of undergraduates. *Learning and Individual Differences*, 45, pp.214-221. https://doi.org/10.1016/j.lindif.2015.12.019

Grün, B. and Zeileis, A. (2009). Automatic Generation of Exams in R. *Journal of Statistical Software*, 29(10), pp.1-14. https://doi.org/10.18637/jss.v029.i10

Hilliam, R., and Calvert, C. (2019). Interactive statistics for a diverse student population. *Open Learning*, 34(2), pp.203-210. https://doi.org/10.1080/02680513.2017.1386095

Jackson, D.C. (2021). Mathematics Support – Essential at the Best of Times, Crucial at the Worst of Times. *International Journal of Innovation in Science and Mathematics Education*, 29(3). https://doi.org/10.30722/IJISME.29.03.003

Koparan, T. (2015). Difficulties in learning and teaching statistics: teacher views. *International Journal of Mathematical Education in Science and Technology*, 46(1), pp.94-104. https://doi.org/10.1080/0020739X.2014.941425

Lalayants, M. (2012). Overcoming Graduate Students' Negative Perceptions of Statistics. *Journal of Teaching in Social Work*, 32(4), pp.356-375. https://doi.org/10.1080/08841233.2012.705259

Lalonde, R.N. and Gardner, R.C. (1993). Statistics as a Second Language? A Model for Predicting Performance in Psychology Students. *Canadian Journal of Behavioural Science*, 25(1), pp.108-125. https://doi.org/10.1037/h0078792

Lear, J.L., Ansorge, C. and Steckelberg, A. (2010). Interactivity/Community Process Model for the Online Education Environment. *Journal of Online Learning and Teaching*, 6(1), p.71.

Liu, D.Y.T., Bartimote-Aufflick, K., Pardo, A., and Bridgeman, A.J. (2017). Data-Driven Personalization of Student Learning Support in Higher Education. In A. Peña-Ayala, ed. *Learning Analytics: Fundaments, Applications, and Trends.* Springer. https://doi.org/10.1007/978-3-319-52977-6_5

Lyons, I.M. and Beilock, S.L. (2012). Mathematics anxiety: separating the math from the anxiety. *Cerebral Cortex*, 22(9), pp.2102-2110. DOI: https://doi.org/10.1093/cercor/bhr289

Macher, D., Paechter, M., Papousek, I. and Ruggeri, K. (2012). Statistics anxiety, trait anxiety, learning behavior, and academic performance. *European Journal of Psychology of Education*, 27(4), pp.483-498. https://doi.org/10.1007/s10212-011-0090-5

Marshall, E.M., Mahmood, B., Alexander, C., Bock, M., Haigney, D., Jack, E. and Verrier, D. (2022). The impact of remote teaching on statistics learning and anxiety, *MSOR Connections*, 20(1), pp.90-101. https://doi.org/10.21100/msor.v20i1.1312

Marshall, E.M., Riach, A., Shaker, A.J. and Haigney, D. (2022). Are students too anxious for statistics anxiety workshops? *MSOR Connections*, 20(2), pp.94-105. https://doi.org/10.21100/msor.v20i2.1287

Marshall, E.M., Rowlett, P., Verrier, D. and Hunt, T. (2021). Can pre-course anxiety and attitudes predict grade? In *MERI research symposium proceedings 2021*, pp.30-31.

Marshall, E.M., Staddon, R.V., Wilson, D.A. and Mann, V.E. (2017). Addressing maths anxiety within the curriculum. *MSOR Connections*, 15(3), p.28. https://doi.org/10.21100/msor.v15i3.555

McArthur, J. (2023). Rethinking authentic assessment: work, well-being, and society. *Higher Education,* 85, pp.85-101. https://doi.org/10.1007/s10734-022-00822-y

Onwuegbuzie, A.J. and Wilson, V.A. (2003). Statistics Anxiety: Nature, etiology, antecedents, effects, and treatments--a comprehensive review of the literature. *Teaching in Higher Education*, 8(2), pp.195-209. https://doi.org/10.1080/1356251032000052447

O'Shea, S., Stone, C. and Delahunty, J. (2015). "I 'feel' like I am at university even though I am online." Exploring how students narrate their engagement with higher education institutions in an online learning environment. *Distance Education*, 36(1), pp.41-58. https://doi.org/10.1080/01587919.2015.1019970

Paechter, M., Macher, D., Martskvishvili, K., Wimmer, S. and Papousek, I. (2017). Mathematics Anxiety and Statistics Anxiety. Shared but Also Unshared Components and Antagonistic Contributions to Performance in Statistics. *Frontiers in Psychology*, 8, pp.1196-1196. https://doi.org/10.3389/fpsyg.2017.01196

Russell, E. (2022). Remote active learning. *MSOR Connections*, 20(1), pp.46-55. https://doi.org/10.21100/msor.v20i1.1308

Shaker, A.J., Hurst, P.S. and Marshall, E.M. (2021). The effect of Kahoot on undergraduate student anxiety and confidence when studying statistics. *MSOR Connections*, 19(2), pp.41-54. https://doi.org/10.21100/msor.v19i2.1245

Thomas, T., Jacobs, D., Hurley, L., Martin, J., Maslyuk, S., Lyall, M. and & Ryan, M. (2019). Students' perspectives of early assessment tasks in their first-year at university. *Assessment & Evaluation in Higher Education*, 44(3), pp.398-414. https://doi.org/10.1080/02602938.2018.1513992

Tishkovskaya, S. and Lancaster, G.A. (2012). Statistical Education in the 21st Century: A Review of Challenges, Teaching Innovations and Strategies for Reform. *Journal of Statistics Education*, 20(2). https://doi.org/10.1080/10691898.2012.11889641

Wilson, V.A. (1999). Reducing statistics anxiety: A ranking of sixteen specific strategies. *Proceedings of annual meeting of the Mid-South Educational Research Association, Point Clear, AL.*

Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: some interesting parallels. *British Journal of Educational Psychology*, 61(3), pp.319-328. https://doi.org/10.1111/j.2044-8279.1991.tb00989.x

CASE STUDY

Development and analysis of a Numbas diagnostic tool for use in a mathematics refresher program

Donald Shearman, Mathematics Education Support Hub, Western Sydney University, Sydney, Australia. Email: d.shearman@westernsydney.edu.au
Shatha Aziz, School of Computer, Data and Mathematical Sciences, Western Sydney University, Sydney, Australia. Email: s.aziz@westernsydney.edu.au
Jim Pettigrew, School of Mathematics and Statistics, University of New South Wales, Sydney, Australia. Email: j.pettigrew@unsw.edu.au

Abstract

We describe the development and analysis of an online diagnostic tool implemented in the Numbas e-learning system and used in an Australian university mathematics refresher program. Following the rapid transition to online delivery of the refresher program caused by COVID-19, the diagnostic instruments and methods used within the pre-pandemic, in-person, version of the program were lost. In 2022, we undertook to revive them in a way that would honour their original diagnostic purpose but offer a more sophisticated approach utilising the Numbas diagnostic exam type. Improvement of the tool after its initial deployment has involved the use of Rasch-based item analysis and recursive refinement of the knowledge map underlying the items.

Keywords: diagnostic tool, mathematics refresher program, knowledge map, item analysis, Numbas.

1. Introduction

For more than 15 years prior to the COVID-19 pandemic, the Mathematics Education Support Hub (MESH) at Western Sydney University (WSU) offered a series of mathematics and statistics refresher workshops for incoming - primarily engineering - students. The workshops were designed to prepare students for the mathematical requirements of their studies, and included content covering topics in basic algebra, trigonometry, statistics, and calculus. Five workshops in these areas (two for algebra) were taught face-to-face over a 3-week period. Students taking the algebra and trigonometry workshops were given pre- and post-tests to measure their skill improvement. The pre-test was also used to place students into equal-ability groups, and so served as a crude diagnostic classifier. All of the lesson and exercise material for each workshop was contained in a printed booklet provided to students at the start of the workshop (this material was also available online during the year for out-of-class learning). These provided a highly structured curriculum for each workshop.

Due to the pandemic, it was necessary to make the workshops available for online study at very short notice. To facilitate a quick transition, we used existing external videos and web resources to present the content of each workshop, along with in-house skill development problems built using the Numbas e-learning system. While the pre- and post-tests were still available online, the pre-test's diagnostic function was largely lost. This limited the capacity of MESH educators to classify students according to skill level, and thus inhibited their work in supporting students' preparation for university study.

2. Plan to develop a diagnostic tool

After three years of online operation, anecdotal evidence indicates that many students find the amount of content covered in the Maths Start workshops overwhelming. They are unable to determine for themselves which selection of lessons and exercises to study in order to improve specific skill deficiencies in topics required for their study. In response, we have developed an online diagnostic tool using Numbas. The tool is designed to identify for each attemptee a selection of topics where mastery criteria have not been met, and hence where further skill development should be focussed.

We have built and deployed diagnostic tools for four of the five Maths Start workshops. We have chosen to refer to our implemented instrument as a diagnostic ‘tool’ rather than ‘test’, as the former term better captures the essential enabling function of the instrument. This paper focuses on the first of these to be built, for the Algebra 1 workshop, in 2023.

3. Diagnostic tool design

The definition of a ‘diagnostic tool’ adopted for this study is an instrument that enables students to identify their areas of skill mastery/non-mastery within a given knowledge domain (Rylands and Shearman, 2022). When considering the design parameters of a suitable diagnostic tool, we were guided by two broad factors: 1. the mechanism by which the tool would identify a students’ areas of mastery/non-mastery; and 2. the amount of time it would take a typical student to complete a full attempt (a measure of the diagnostic efficiency of the tool).

A form of computer adaptive test was chosen to meet these criteria. This type of test is designed to update a latent measure of the attemptee’s ability while they are answering questions. If they answer a question correctly, they are only presented with more difficult questions in the hierarchically-ordered set of questions to which it belongs (or taken to another set if the most difficult question in the current set has been mastered). Conversely, if they answer a question incorrectly, they are stepped through easier questions in the hierarchy until mastery is met or the questions in the set are exhausted. Attempt times for these tests are usually shorter than for traditional tests because attemptees only complete a subset of the full question set. In tailoring the questions presented to the individual attemptee according to difficulty, computer adaptive tests aim to find the attemptee’s threshold skill level and give a more accurate assessment of their knowledge (Meijer and Nering, 1999).

There are many subvariants of computer adaptive test designs; the one chosen for this study relies on an underlying ‘knowledge map’ consisting of hierarchically-arranged concepts and associated topics. Throughout all of the development and analysis described in this case study, we have assumed that knowledge of a concept is demonstrated by way of mastery of a specific skill (or skills). In practical terms, the implemented design is such that:

- if a question is answered correctly, it is assumed that any question about a concept on which that question relies would also be answered correctly, so these questions are marked as correct;
- if a question is answered incorrectly, it is assumed that any question about a concept which relies on the concept being assessed would also be incorrect, so these questions are marked incorrect.

In developing the diagnostic tool for this study, the first stage was the construction of the abovementioned knowledge map, which acted as a formal network of the necessary connections between the concepts and associated topics being assessed.

4. The knowledge map

The knowledge map used in this study was developed by reviewing the sections and subsections of the Algebra 1 workshop and making links between those exhibiting clear and logical conceptual connections. The result, a directed acyclic graph (DAG), was reviewed by other members of the MESH team and opinions on the included topics (nodes) and their hierarchical connections (directed edges) were discussed and resolved. A section of the final map for the Algebra 1 content is shown in figure 1.

This was created using the graph visualisation application Gephi. Note that only the numbered items are topics; the unnumbered items are learning objectives, included in the map for convenience (as they show the grouping of topics by learning objective).

Once the construction of the knowledge map was complete, the next task was to assign a question to each of its nodes.

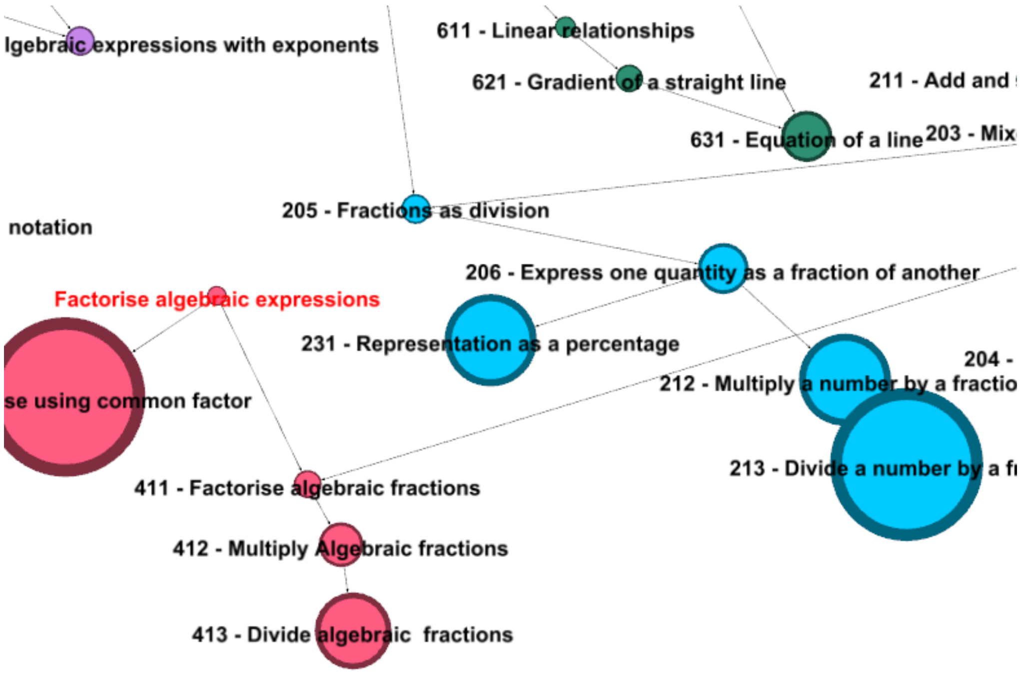


Figure 1. A section of the knowledge map for the Algebra 1 workshop.

5. Implementing the diagnostic tool using Numbas

The system used to implement the Algebra 1 diagnostic tool is the Numbas diagnostic exam (more precisely, the diagnostic mode of the generic Numbas exam object). At the heart of this system is an algorithm, known as DIAGNOSYS, that acts on questions whose underlying concepts are grouped within ‘learning objectives’. Individual concepts within learning objective groups are classed as ‘topics’. For example, in the Algebra 1 diagnostic tool, we have defined ‘Factorise algebraic expressions’ as a learning objective and ‘Factorise algebraic fractions’ and ‘Multiply algebraic fractions’ as topics within this learning objective (see figure 1). The system allows each topic to be assigned a question, hierarchically linked to other topics, and grouped within a specific learning objective.

DIAGNOSYS conforms to the general principles of computer adaptive test design: the ‘next’ question presented to the attemptee (except the first) depends on their responses to previous questions; no questions are presented whose underlying concepts are assumed by the algorithm to have been mastered or unable to be mastered. The effect of the hierarchical arrangement of questions in a Numbas diagnostic exam is that:

- an incorrect answer to a question causes the system to mark all harder questions on the same hierarchical path as incorrect;
- a correct answer to a question causes the system to mark all easier questions on the same hierarchical path as correct.

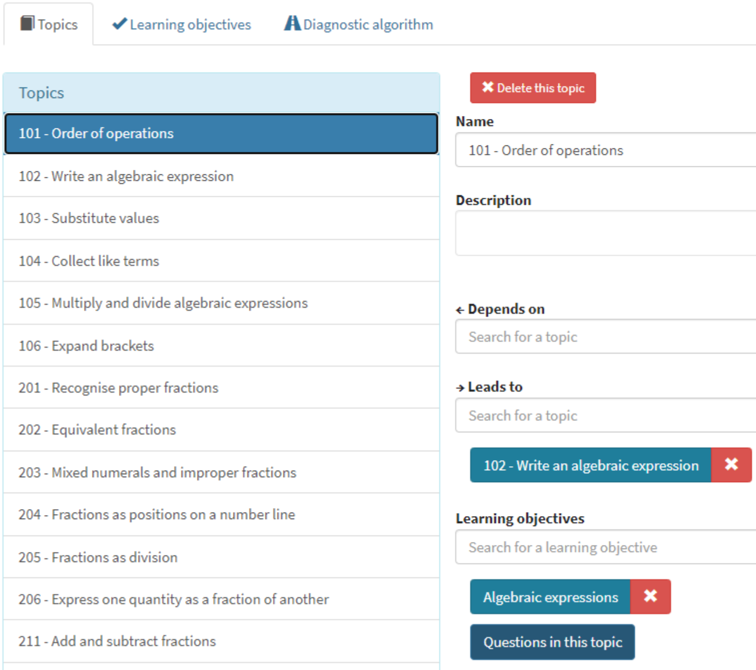


Figure 2. The Numbas authoring interface for topic linking.

In practice, after almost 12 months of use, this has reduced attempt times for the Algebra 1 diagnostic tool. We believe this has led to higher completion rates, and, ultimately, enhanced MESH's ability to support incoming students unsure of how to tackle their skill refreshment.

The knowledge map was written into DIAGNOSYS by creating topics grouped within learning areas and the necessary connections between these topics. Topics were paired by using either the 'leads to' or 'depends on' directed connections (see figure 2). This task was complex as extra care was needed to ensure faithful translation of the knowledge map and accurate reconstruction of the DAG in Numbas.

To complete the construction of the tool, each topic in Numbas was assigned a question. This process was simplified by the fact that the Algebra 1 curriculum, whose structure has been refined over two decades by educators within MESH, coexists with a set of Numbas questions (used in non-diagnostic quizzes prior to the development of the diagnostic tool). For each topic, a search was conducted for a suitable existing question. If one was found, it was either adopted unchanged or modified. In exceptional cases, new questions were created in the absence of suitable existing candidates.

In the second half of 2022, regular review meetings addressed question suitability and the composition of the knowledge map (whose development was recursive). Other educators within MESH who did not participate in these review meetings trialled early versions of the Algebra 1 tool; their feedback allowed for corrections and refinements. The process of review and refinement was repeated until January 2023, when we deployed the final version of the tool on WSU's learning management system.

6. Deployment

The Algebra 1 diagnostic tool was loaded onto a dedicated Learning Tools Interoperability (LTI) server and made available to all WSU students on the Maths Start site. The tool is still 'live' and as of mid October 2023 has been attempted over 800 times by 599 unique students. The first attempt was on 1 February 2023.

7. Analysis

Following deployment of the diagnostic tool in 2023, an analysis of the response data was conducted to determine whether the questions were functioning according to their design specifications. Particular attention was given to the relationship between the questions and the knowledge map, and the extent to which the logical connections between questions were validated by the patterns in the response data.

The analysis method was to compare two sets of question responses (hereafter, we'll refer to questions interchangeably as items): 'raw' responses for scored student attempts, and 'implied' responses for unattempted items whose scores were implied by DIAGNOSYS and its application of the knowledge map. In each case, correct and incorrect responses were assigned a score of 1 and 0 respectively. The implied scoring works as follows: if an attemptee answers an item correctly (call it item A, for instance) and items B, C and D are set as easier than item A, then the tool assumes the attemptee will also answer items B, C and D correctly - and therefore not present these items to them. Figure 3 gives an example of these relationships for four connected items in the knowledge map.

The knowledge map was built in Gephi such that items (nodes) connected (via edges) to other items but lying below them in the graph were dependent: the knowledge and skills required to correctly answer item A (below) include and extend those required to correctly answer items B, C and D (above). Thus the implied scoring algorithm worked by assigning a score of 1 to all easier items above a correctly-answered item and connected to it in the knowledge map (the 1s 'float'), and assigning a score of 0 to all harder items below an incorrectly-answered item and connected to it in the knowledge map (the 0s 'sink').

The response data was exported from the Numbas LTI server as a JavaScript Object Notation (JSON) file and imported into the R statistical analysis program using the 'jsonlite' package. The knowledge map was exported from Gephi as a Graph Exchange XML Format (GEFX) file for manipulation in R.

The attempts for which each of the 50 items received an implied score of 0 or 1 were collected as the implied-scored response data. There were 412 such attempts, giving a 412 by 50 data frame in R. Their associated raw-scored response data was stored in an identically-dimensioned data frame, populated by 0s and 1s as well as NAs for the missing scores (see figure 4).

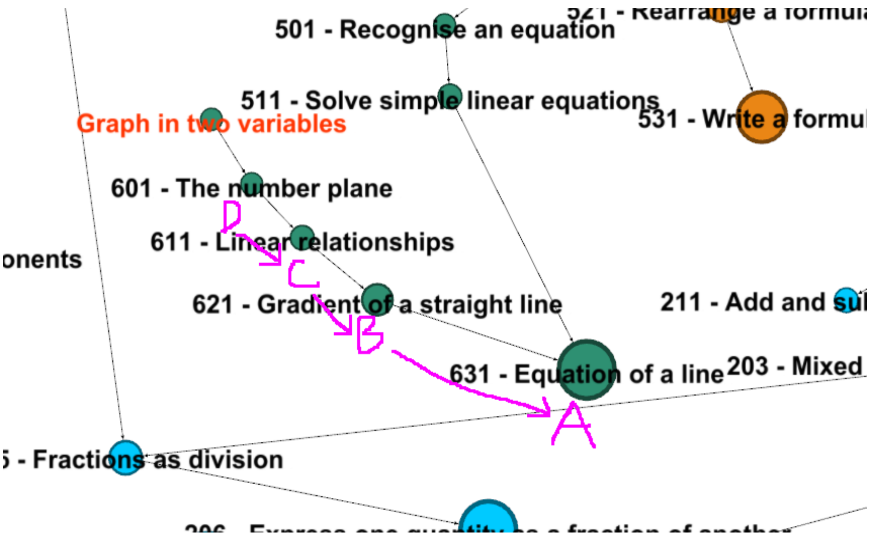


Figure 3. Graphical relationship between items A, B, C and D. Item D is the easiest of the four, C the next easiest, and so on. A is the hardest item in the set.

	Q1_score	Q2_score	Q3_score
1	NA	0	0
2	NA	0	0
3	NA	NA	NA
4	NA	NA	NA
5	NA	1	1
6	NA	NA	NA
7	NA	0	NA
8	NA	0	0
9	NA	1	0

	Q1_implied_score	Q2_implied_score	Q3_implied_score
1	1	0	0
2	1	0	0
3	1	1	1
4	1	1	1
5	1	1	1
6	1	1	1
7	1	0	1
8	1	0	0
9	1	1	0

Figure 4. Snippets of the raw- and implied-scored data frames used in R.

Rasch modelling was applied to the raw-scored and implied-scored response data separately to determine, first, item difficulties for the items in each case and then other properties such as item discrimination (point-biserial correlation), and various statistical measures of item and model fit (see Wu, Tam and Jen, 2016). It should be noted that this use of Rasch modelling is unconventional for at least two reasons: 1. the dichotomously-scored data is incomplete in that missing and implied scores represent 'responses' to items that were not presented to students; and 2. randomisation of variables in the tool meant that different students attempted (slightly) modified versions of the same items. The authors make no claim as to a rigorous, scientifically-robust use of this modelling, but note that it has provided an alternative methodology to Classical Test Theory in assigning measures of difficulty to each of the 50 raw-scored and implied-scored items. Furthermore, the ensuing comparison of item rankings by difficulty (raw-scored versus implied-scored) drew attention to a number of potentially dysfunctional items that, upon close inspection, were revealed to be so.

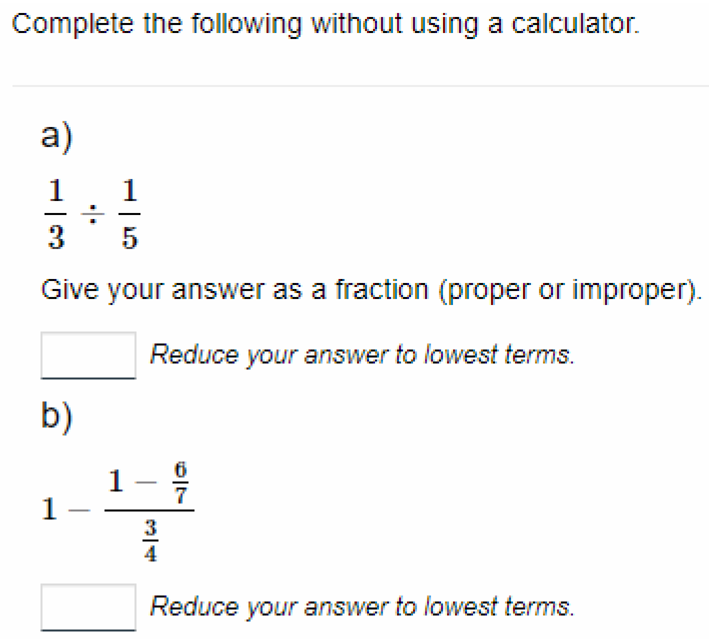
The raw-scored and implied-scored items were ranked (separately) from least to most difficult. For example, the three least difficult raw-scored items were questions 22, 24 and 25, while the three least difficult implied-scored items were questions 7, 1 and 16. The unsigned difference in ranking was calculated for each item, and this yielded a list of 10 items where the difference was greater than or equal to 15. (Most rankings were close, for example 35 of the 50 items had an unsigned difference of 8 or less.) Our analysis proceeded by assuming that these items were anomalous in the sense that they were malfunctioning according to their design, or irregular due to unexpected attemptee behaviour or item interaction.

The final phase of the analysis was to use the results of the Rasch modelling to guide a critical inspection of the anomalous items. We shall present two examples of the findings of this work.

7.1 Question 11 (raw-scored rank: 39; implied-scored rank: 5)

This question is shown in figure 5. It aims to assess skill in dividing one fraction by another. According to the raw-scored responses, it was ranked as the 12th hardest item; but according to the implied-scored responses, it was ranked as the 5th easiest (giving a rank difference of 34). What could explain this difference?

We believe that the answer relates to part (b) of the question, mastery of which would require arithmetic skill beyond the direct division assessed in part (a). As figure 6 illustrates, the attemptees who got this question - appearing as ‘205 - fractions as division’ - wrong would not have been presented with any of the questions below it on the hierarchical path (note that 56 of the 412 attempts were in this category). This means that they would have received the implied score of 0 for all of these ‘harder’ questions (recalling that the 0s ‘sink’). In particular, an incorrect raw response to Question 11 would result in an implied incorrect response to Question 15 (identified as ‘213 - Divide a number by a fraction’). Conversely, a correct raw response to Question 15 would result in correct implied responses for all of the ‘easier’ questions above it in the path (recalling that the 1s ‘float’).



Complete the following without using a calculator.

a)

$$\frac{1}{3} \div \frac{1}{5}$$

Give your answer as a fraction (proper or improper).

Reduce your answer to lowest terms.

b)

$$1 - \frac{1 - \frac{6}{7}}{\frac{3}{4}}$$

Reduce your answer to lowest terms.

Figure 5. Question 11 of the diagnostic tool (identified in the knowledge map as ‘205 - Fractions as division’).

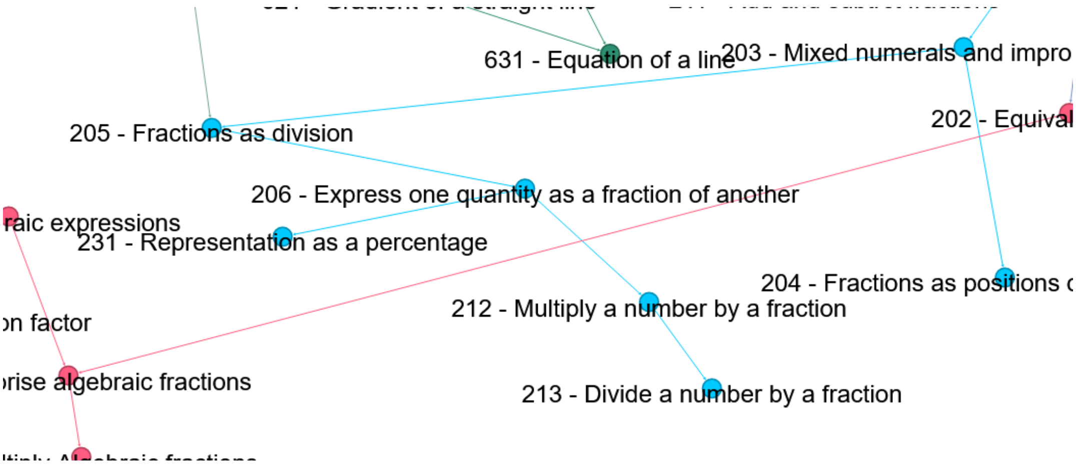


Figure 6. A snippet of the knowledge map, showing the location of Question 11 (identified as ‘205 - Fractions as division’), and its connection to harder questions below and easier questions above it.

Question 15 is shown in figure 7. Is it reasonable to assume that attemptees who answered this question correctly would be able to answer Question 11(b) correctly? We believe not and so have adjusted Question 11 to better align it with its underlying concept and location in the knowledge map.

If the above analysis is correct, then the reason Question 11 was ranked as significantly easier under implied scoring compared to raw scoring is that many attemptees (167 of 412) were presented with Question 15, answered it correctly and subsequently were not presented with any of the questions higher than it on the path (in particular Question 11 - see figure 6). All of these higher items, assumed easier, would have been given an implied score of 1.

Complete the following without using a calculator.

a)

$12 \div \frac{1}{8}$

b)

$\frac{3}{7} \div 5$

Reduce your answer to lowest terms.

Figure 7. Question 15 of the diagnostic tool (identified in the knowledge map as ‘213 - Divide a number by a fraction’).

Version 1

Libby wants to write the cost formula which represents the cost of producing paper clips. The cost function C equals the initial cost added to 5 times the marginal cost. Let F, M represent the initial and marginal costs respectively. Can you help Libby to write the cost fomula?

$C =$

Version 2

Let b, h be the base length and height of a triangle respectively. The area A of this triangle equals half the base by the height. Write the formula which represents the area of the triangle.

$A =$

Figure 8. Both versions of Question 46 of the diagnostic tool (identified in the knowledge map as ‘531 - Write a formula’).

7.2 Question 46 (raw-scored rank: 15; implied-scored rank: 42)

This question is shown in figure 8 (both versions; noting that either is presented to attemptees at random). It aims to assess skill in translating a worded description of a formula to its symbolic, mathematically-notated form. According to the raw-scored responses, it was ranked as the 15th easiest item; but according to the implied-scored responses, it was ranked as the 9th hardest (giving a rank difference of 27).

We propose that the reason Question 46 (appearing as '531 - Write a formula' in the knowledge map - see figure 9) was ranked as significantly harder under implied scoring compared to raw scoring is that many attemptees would have been presented with Question 2 (appearing as '102 - Write algebraic expressions', see figures 9 and 10), answered it incorrectly and subsequently be given a mark of 0 for all of the 'harder' questions below it in the hierarchical path (in particular, Question 46). This means that they would have received the implied score of 0 for Question 46 (assumed to be harder). But in our judgement, version 2 of Question 46 is easier then Question 2, and this is reflected in a disproportionately high number of correct raw-scored responses.

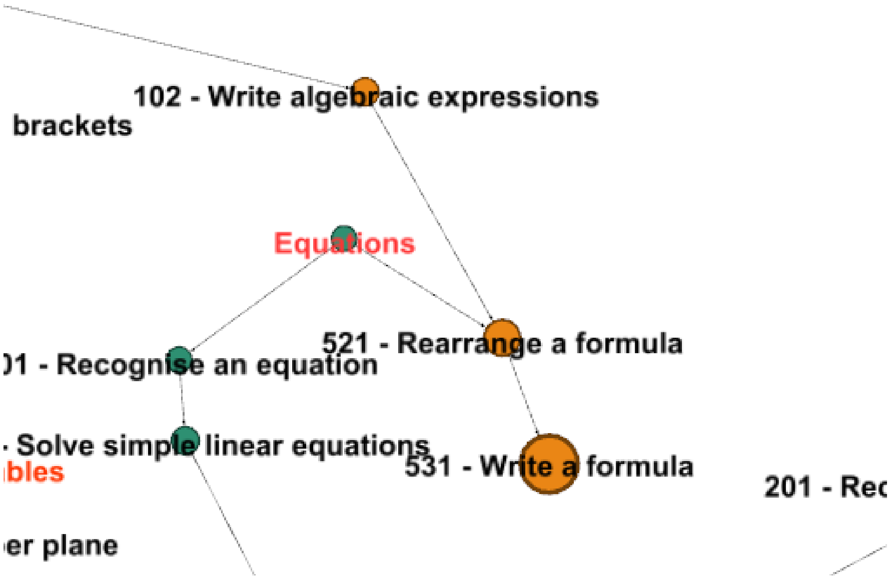


Figure 9. A snippet of the knowledge map showing the location of Question 46 (identified as '531 - Write a formula'), and its connection to easier questions above it.

H is used to represent a standard hourly rate of pay in dollars.

Which one of the following expressions would represent the amount of pay earned for 10 hours work, where 3 of those hours were paid at one and a half times the standard rate?

○ $13H$ ○ $11.5H$ ○ $\frac{10}{1.5}H$ ○ $14.5H$

Figure 10. Question 2 of the diagnostic tool (identified in the knowledge map as '102 - Write algebraic expressions').

8. Conclusion

In this case study we have described the process by which we developed and analysed an online diagnostic tool for deployment within a university mathematics refresher program. The tool's diagnostic power is derived from a knowledge map defining and connecting key concepts in a basic algebra knowledge domain. Use of the tool has enabled attemptees to determine their level in this knowledge domain in a manner that is more efficient and personalised than non-adaptive alternatives (featuring banks of multiple-choice questions, for example). From the educator's perspective, the tool has allowed for the optimised delivery of support services and learning resources that are tailored to the individual attemptee based on their location in the knowledge domain. An important difference between our tool and other computer-adaptive diagnostic systems is that 'next question' selection is based on the attemptee's trajectory through a knowledge map rather than a real-time estimate of their ability (matched with questions whose difficulty is on the same logit scale in the case where Rasch and related analysis methods are used).

The tool has been built using the Numbas e-learning system, whose essential DIAGNOSYS algorithm operates according to principles of computer adaptive test design. The refresher program's existing curriculum provided an unrefined hierarchical structure for the mathematical concepts and associated skills covered in the program, and this led to the creation of the abovementioned knowledge map defining their necessary connections. The tool was built using a Numbas exam object (diagnostic mode), which was designed carefully to honour the logical connections in the knowledge map. A combination of new and existing Numbas questions was used to populate the exam. Analysis of the item responses for 826 attempts of the tool has revealed significant differences in the difficulty ranking of some raw-scored and implied-scored questions, and this drew attention to anomalies in the design of a subset of questions and their hierarchical relationships in the knowledge map. Addressing these anomalies has involved modifying questions and reconfiguring the knowledge map. In this way, improvement of the tool has been recursive.

Though it has not been developed as an instrument for formative or summative assessment per se, the design of the diagnostic tool would allow for ready applications in these ways. The Numbas system allows for exams to be switched from diagnostic to mastery mode, enabling attemptees to hone their skill with unlimited question attempts. An adapted version of the tool could also be used for delivery of summative assessments. Summative computer adaptive tests are used in many of Microsoft and CISCO's Computer Certification Tests, for example. In broad terms, the idea here is that the test-taker's response patterns allow the system to find their 'point of equilibrium' in settling on the questions – hence concepts – positioned at their level of ability.

An extension of this study would be to compare the existing 'subjectively constructed' knowledge map to one that has been generated using an 'objective' statistical method. The latter would involve administering all of the tool's questions to a sample of students and using one of many (typically naïve Bayesian) approaches to define a network of connections between the concepts underlying the questions. (There are many examples of such approaches in the literature.) We could then compare the 'objective' and 'subjective' knowledge maps and if the differences are minimal adopt the latter as valid for the constrained diagnostic purposes of the tool. Another extension would be to look across the knowledge map and calibrate topics based on the Rasch-determined difficulty of their associated questions. This would add a 'depth' dimension to the knowledge map in introducing a vertical scale to its nodes and edges (some edges being longer than others where the Rasch analysis has determined the difference in the difficulty of the topics they connect is greater, for example).

9. References

Meijer, R. R. and Nering, M. L. (1999). Computerized Adaptive Testing: Overview and Introduction. *Applied Psychological Measurement*, *23*(3), pp.187-194. https://doi.org/10.1177/01466219922031310

Rylands L. and Shearman D. (2022). Diagnostic tests: Purposes and two case studies. *MSOR Connections*, 20(3), pp.45-54. https://doi.org/10.21100/msor.v20i3.1281

Wu, M., Tam, H.P. and Jen, T.H. (2016). Educational measurement for applied researchers. *Theory into practice*. Springer Singapore. https://doi.org/10.1007/978-981-10-3302-5

CASE STUDY

Engaging students via interactive lecture activities inspired by common classroom practice

Cerian Brewer, School of Mathematical Sciences, University of Nottingham, Nottingham, UK.
Email: cerian.brewer1@nottingham.ac.uk

Abstract

This article summarises the author's experience transitioning from Further Education (FE) to Higher Education (HE) teaching in 2022-23. It identifies areas for concern in the author's setting and outlines the author's strategies for overcoming these obstacles, utilising their background in FE. In particular, this article explores common, engaging classroom activities and outlines why these activities require adapting for use in HE. Inspired by these classroom activities, this article presents a selection of interactive lecture activities promoting student engagement via formative assessment, adaptive teaching, and student ownership. It summarises how the activities have evolved over time and reflects on their use. Finally, the author reviews the success of the activities, discusses students' reactions to them, and outlines future plans.

Keywords: engagement, interactivity, formative assessment, adaptive teaching, student ownership.

1. Background

The COVID-19 pandemic has had a massive effect on education over the last few years. Anecdotally, it is common to hear colleagues discuss the effect of the pandemic on student attendance and engagement, with lecture theatres sometimes containing only a handful of students, and students being reluctant to engage in activities: this is demotivating for staff and students alike, and a common concern in the author's setting.

The author has predominantly been engaged in service teaching since joining the HE sector, teaching maths for engineering students. This comes with its own difficulties since these students did not choose to study maths at university: peers highlight attendance and engagement as particular areas for concern when teaching engineering students.

Whilst preparing teaching materials, the author has targeted student engagement in the hope that this might also lead to improved student attendance. The aim has been to promote student engagement via interactive lecture activities, based on common, engaging classroom strategies, with a particular focus on formative assessment, adaptive teaching, and student ownership. The author has also aspired to create active rather than passive learners, and to re-ignite students' enthusiasm for learning in person and encourage students back to the lecture theatre.

The benefits of interactive lectures are well established and have been a topic of discussion for many years. Interactive lectures allow more time to think than passive lectures, allowing students to determine their own understanding and to ask questions (Rodger, 1995). Mannison et al. (1994) report that students' involvement, morale, and interest in the lecture material is positively impacted by the inclusion of interactive teaching strategies, and Rodger (1995) finds that students "*overwhelmingly prefer*" interactive lectures, for example commenting "*The interactive format was very helpful and made the class more interesting*."

In more recent years, Robson et al. (2022) report that lectures have become more interactive since the pandemic, and participants in their study recognise the importance and value of interactive teaching, indicating that lectures in a post-pandemic world should make greater use of interactivity. In fact, Robson et al. (2022) report that most staff would not return to conventional lecturing post-pandemic, instead favouring more interactive lectures.

This article focuses on the development and implementation of a selection of interactive lecture activities promoting student engagement in a level 2 maths module for around 80 engineering students. Section 2 reviews a selection of pedagogical tools that the author aims to incorporate into the activities: in particular, formative assessment, adaptive teaching, and student ownership. Section 3 introduces common, engaging classroom activities that promote student engagement via the aforementioned tools and outlines why they require adapting for use in HE. Section 4 presents the interactive lecture activities that have been inspired by these common, engaging classroom activities, and summarises how they have evolved over time and provides general reflections on their use. Finally, section 5 reviews the success of the activities, discusses students' reactions to them, and outlines the author's future plans.

2. Overview of pedagogical tools

The benefits of formative assessment are widely acknowledged throughout the education sector. Black and Wiliam (2010) argue that formative assessment should be at the heart of effective teaching, and Stiggins and DuFour (2009) describe formative assessment as "*one of the most powerful instructional tools available to a teacher*". This article considers the use of formative assessment, embedded in a set of interactive lecture activities, as a tool for boosting student engagement, which Stiggins (2002) claims is more meaningful when teachers embrace the idea of assessment *for* learning, rather than assessment *of* learning.

To be most effective, Stiggins and DuFour (2009) advocate the use of formative assessment continuously, whilst learning is happening, rather than using it periodically. As such, the activities presented in this work have been used regularly to support day-to-day learning. In the specific context of teaching maths for engineering, formative assessment is also encouraged by Sazhin (1998).

If used properly, the evidence gained from formative assessment should be used to adapt teaching (Black and Wiliam, 2010). A teacher's ability to adapt is another key component of effective teaching according to Darling-Hammond and Bransford (2007) and Hattie (2008). So-called adaptive teaching describes a teacher's unplanned response to a stimulus, for example a student contribution or misconception, which deviates from the lesson plan (Hardy et al., 2019). Adaptive teaching requires the teacher to monitor students' understanding and classroom proceedings and use this monitoring, along with subject and pedagogical knowledge, to alter instruction to best support students' needs (Duffy, 2006, Vaughn and Parsons, 2013). These alterations should focus on meeting the identified needs of students whilst retaining high expectations for their learning (Hardy et al., 2019). Adaptive teaching has also been described as an important factor in ensuring equal opportunities for students to achieve their goals (Hardy et al., 2019). The activities presented in this work have therefore been designed to promote adaptive teaching and to allow for flexibility.

Formative assessment can also contribute to student ownership of learning (Bloom, 1984) and the promotion of student ownership via formative assessment strategies has been shown to result in high student engagement (Brookhart et al., 2009). Students can feel empowered when they take an active role in their education and become meaningfully engaged in their learning: taking ownership in this way can promote goal setting, self-assessment, and self-determination (Chan et al., 2014). As such, the activities presented in this work have been designed to encourage student ownership of learning.

3. Overview of classroom activities

This section introduces the common, engaging classroom activities that have influenced the author, and highlights reasons why they may need adapting for use in HE. It summarises links to formative assessment, adaptive teaching, and student ownership also.

3.1 Questioning with mini whiteboards

A common classroom activity promoting formative assessment is questioning with mini whiteboards. Swan (2005) recommends this strategy and refers to mini whiteboards as an "*indispensable resource*". The key reasons for this endorsement are summarised below.

When questioning with mini whiteboards is implemented, students can simultaneously present their responses to the teacher and/or to each other, promoting a 'no hands' approach in which the teacher doesn't need to rely on the same students putting their hands up to volunteer answers. It is possible to see at a glance what every student is thinking when they hold up their ideas, and a variety of responses can be collected, for example written or drawn solutions, which can allow teachers to ask new kinds of questions, often beginning "*Show me*..." (Swan, 2005). Overall, it is a quick and effective way to check students' understanding.

As well as promoting formative assessment, this strategy can encourage adaptive teaching since the teacher's instruction can be altered based on students' responses. This gives students an active role to play in their learning which can, in turn, support student ownership.

This strategy needs adapting for HE, however, due to the (potentially very large) number of students present and the (correspondingly large) size of the teaching rooms, not to mention the logistics of distributing mini whiteboards and pens to students. What is perfectly feasible for 20-30 students in a classroom environment becomes an impractical endeavour in an HE context. Section 4 discusses the interactive lecture activities, inspired by questioning with mini whiteboards, that have been developed with these issues in mind.

3.2 Student reactions

The term student reactions is used here to cover a broad range of informal teaching strategies that are used to gauge students' understanding and progress. As such, they can act as formative assessment tools as well as being used to judge the pace of lessons.

A common classroom activity for assessing students' understanding of a problem, after attempting it, is asking students to rate their understanding with a thumb up or down (or somewhere in between). Similarly, asking students to rate their confidence before attempting a problem can provide useful feedback for the teacher. Other strategies for checking confidence/understanding include having one-to-one conversations with students and questioning the class; students asking questions can also give a good indication of their understanding. Picking up on visual cues is also helpful, for example students might look confused, in which case the teacher can provide support.

If students find a certain concept or example difficult, the teacher can choose to spend longer going through it or provide additional scaffolding (Wood et al., 1976). Conversely, if students demonstrate understanding of an example or concept, the teacher can move onto more challenging problems. Using student reactions as a formative assessment technique therefore allows for adaptive teaching and the teacher can alter the pace of the lesson based on students' reactions. Other visual cues, such as students putting their pens down after finishing an example, can also feed into this.

Judging student reactions in a classroom environment is achievable but, due to the (potentially very large) student numbers and room sizes, this strategy needs adapting for HE. Section 4 outlines the activities that have been designed with student reactions in mind, for specific use in HE.

3.3 Worked examples

Worked examples are a popular instructional device which provide an expert's solution for a student to study and emulate; typically, a worked example includes a statement of a problem followed by a model solution, usually presented in a step-by-step fashion (Atkinson et al., 2000). In terms of cognitive load theory (Sweller, 1988), worked examples are a valuable tool since they reduce the cognitive load during learning (Atkinson et al., 2000).

Trafton and Reiser (1993) claim that the most efficient method of study is to present a worked example immediately followed by a similar problem for students to solve, namely a worked example-problem pair. In this case, students use knowledge gained from the worked example and apply it to solving a new problem. Having access to a model solution can also act as scaffolding (Wood et al., 1976) for students.

Similar methods are adopted by Sweller and Cooper (1985), who find that students studying worked example-problem pairs perform significantly better than those simply solving problems, and more recently by Atkinson et al. (2000), who find that lessons incorporating worked example-problem pairs produce better outcomes than lessons presenting a series of worked examples followed by a series of practice problems.

Worked example-problem pairs are a common classroom activity and lend themselves well to adaptive teaching: whilst presenting a worked example, or whilst students work through a similar problem, the teacher's instruction can be altered based on students' understanding/misconceptions etc. Allowing time for students to attempt problems themselves can also promote student ownership since students take an active role in their learning.

Although worked example-problem pairs lend themselves well to an HE setting, due to (potentially very large) student numbers and room sizes, it is difficult to check students' understanding of worked example-problem pairs and/or to provide appropriate scaffolding. This is perfectly manageable for 20-30 students in a classroom using techniques discussed previously, for example asking students to rate their understanding with a thumb up/down/somewhere in between, or simply picking up on visual cues. Section 4 describes the activities that have been designed to allow for understanding checks and scaffolding whilst using worked example-problem pairs in an HE setting.

4. Overview of activities adapted for HE

This section presents the interactive lecture activities, inspired by the common, engaging classroom activities described in section 3, that have been designed to promote student engagement via formative assessment, adaptive teaching, and student ownership. It also summarises how the activities have evolved over time, based on observations of student engagement, and reflects on their use in the author's setting.

4.1 The activities

The following lecture activities have been developed using the audience response tool Mentimeter (https://www.mentimeter.com/), which allows for a variety of polling techniques, including answering multiple choice questions, providing open ended responses, rating statements, and ranking items. Participants join a Mentimeter presentation on their own device and can advance either at the presenter's pace or at their own pace, based on the settings chosen by the creator. Responses are all anonymous and, as well as the polling options listed above, participants can interact with slides via a range of reactions, for example thumbs up or thumbs down. The author has not explored all polling options but, in this section, outlines the key strategies used in an attempt to promote student engagement whilst teaching around 80 students enrolled on a level 2 maths module for engineering students. The author also notes that the use of the free version of Mentimeter, as with any other audience response tool, has its limitations.

Russell (2022) notes that online polling can create opportunities for instructor-student feedback, which can happen in real time. In a similar vein, online polling allows for formative assessment and adaptive teaching and can provide similar information to the teacher as questioning using mini whiteboards (see section 3.1). In this section, the author also explores the use of Mentimeter to gauge student reactions (see section 3.2) and to assess student confidence and understanding when implementing worked example-problem pairs (see section 3.3).

Perhaps the most obvious mode of interaction, multiple choice questions are a popular tool for checking students' understanding in a lecture. The author has implemented traditional multiple choice questions using Mentimeter to judge students' understanding of mathematical concepts, but has also used multiple choice questions to evaluate students' enjoyment of subjects. For example, figure 1 shows the multiple choice activity used in the first module lecture, in which students were asked to complete the sentence "*Maths is*…". Results from the activity did not directly change lecture delivery but provided the author with a useful sense of students' prior experience and enthusiasm for the subject.

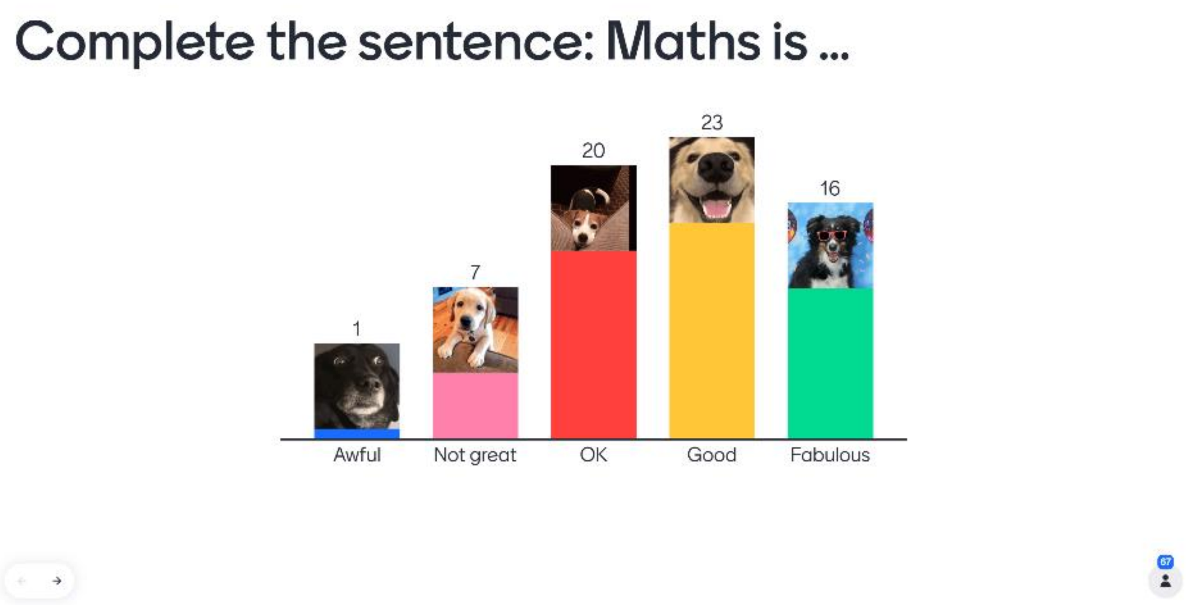


Figure 1. An example of a multiple choice activity created using Mentimeter, used in the introductory lecture to gauge students' feelings towards maths.

The student reactions discussed in section 3.2 have been replicated using Mentimeter's reactions feature, for example students have been asked to react to a slide with a thumb up or a thumb down. This basic feature has proved very useful and has been used in a variety of ways to obtain different information from students. For example, students have been tasked with solving a problem and asked to react to a Mentimeter slide with a thumb up when they have completed the problem or a thumb down if they are unsure where to start. This has created opportunities for adaptive teaching and scaffolding when students react with a thumb down. For example, figure 2 shows the final question from an activity recapping trigonometry, in which students were asked to come up with an expression for $\cos(n\pi)$ in terms of an integer n. The author hoped that students would spot that it alternates between -1 and $+1$ for odd and even values of n, respectively, and start to formulate a mathematical expression for this, i.e. $(-1)^n$. After a few minutes of discussion, several students reacted to the slide with a thumb down, indicating that they were unsure how to approach the problem. The author then scaffolded the problem, sketching the graph of $y = \cos(x)$ and facilitating a class discussion. When such activities have been used, solutions have been displayed or students asked to vote for the answer when most students have indicated that they have completed the task, potentially after scaffolding.

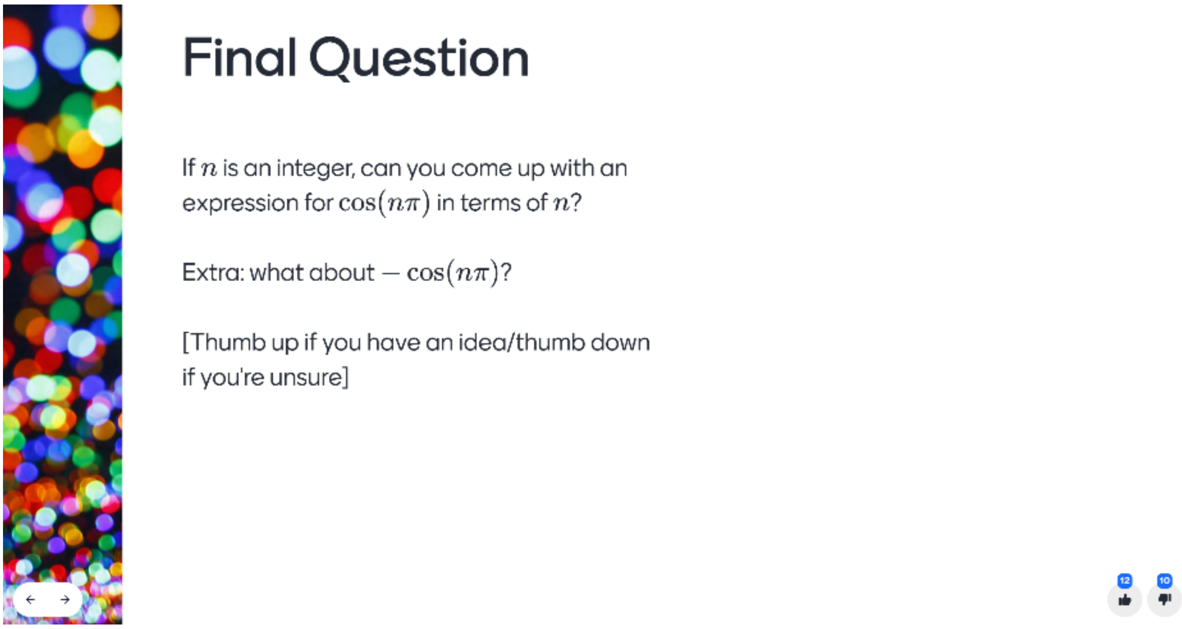


Figure 2. An example of an activity utilising the reactions feature on Mentimeter to gauge students' progress and understanding.

Alternatively, after students have attempted a problem, solutions have been displayed on Mentimeter and students asked to react with a thumb up if they agree with/understand the solution or a thumb down if they disagree with/don't understand it. For example, figure 3 shows the slide following the question displayed in figure 2, in which the solution to the problem was provided. Students reacted positively to this solution slide following the scaffolded class discussion, reassuring the author that the solution was widely understood.

The strategies outlined above have been implemented during worked example-problem pairs: in particular to check students' understanding of the similar problem after attempting it. To gauge students' understanding of the worked example itself, students have also been asked to rate their understanding of a worked example straight after working through it. Similarly, students have been asked to rate their confidence before attempting a similar problem for themselves, after the similar problem has been displayed on the board. Both rating strategies have encouraged adaptive teaching and scaffolding when understanding or confidence have received low ratings.

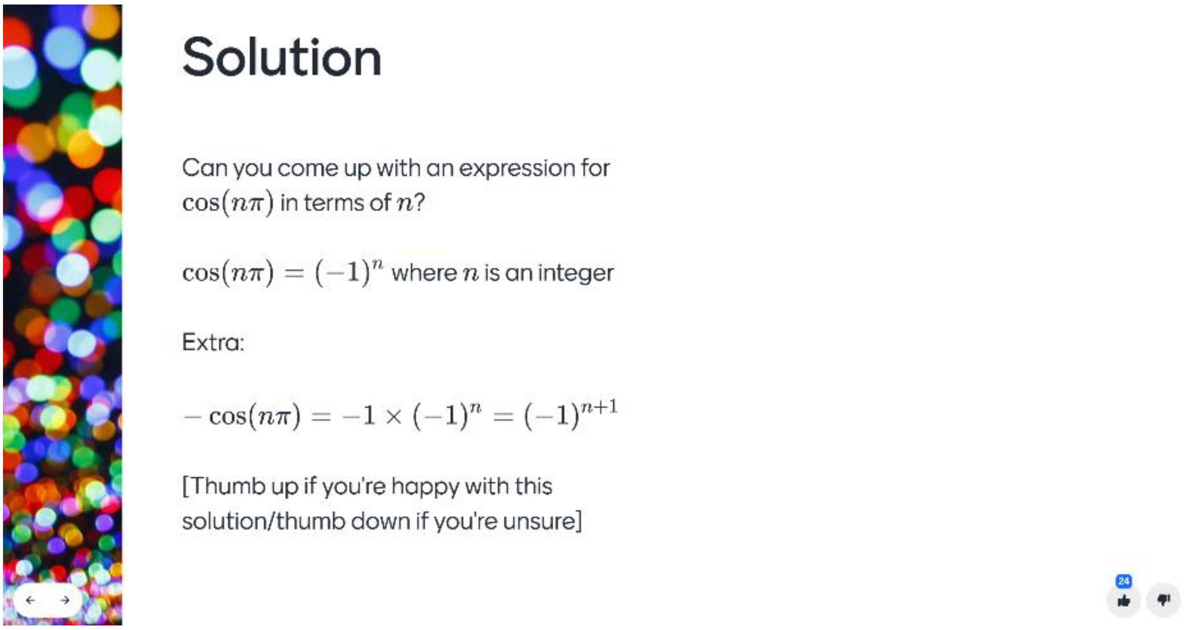


Figure 3. Another example of an activity utilising the reactions feature on Mentimeter to gauge students' understanding. This slide followed on directly from the one displayed in figure 2.

The aforementioned strategies have been used in isolation as quick checks of confidence/understanding, but also combined to create longer activities. In fact, many of the interactive lecture activities used by the author incorporate several of the features discussed above. For example, figures 4-6 display the slides from an activity used in teaching multiple integration. The activity followed a straightforward worked example introducing the topic and started with students rating their confidence before attempting a similar problem. Students then attempted the similar problem, reacting to the second slide to indicate that they had completed it, before voting for the answer on the third slide. The author anticipated fairly high confidence levels and success rates due to the low difficulty of the question and this proved to be a reasonable prediction in this case; if provided with alternative evidence, however, scaffolding would have been provided.

4.2 Adaptations and reflections

The activities outlined in section 4.1 have been updated and adapted over time based on observations of student engagement. For example, early activities provided little in the way of stretch and challenge: students who completed examples quickly sat waiting for their peers to finish and were therefore not engaged for parts of the lecture. As such, extension questions have been included in later activities, where appropriate. For example, figure 7 shows an activity in which a second, more challenging, problem was included for students to try if they worked through the first problem quickly. Longer activities have also been trialled, in which students worked through several questions at their own pace, and reacted to slides along the way to indicate completion/understanding. In accordance with Bloom's taxonomy (Bloom, 1956), these questions were ordered by increasing levels of difficulty, allowing students to access more challenging questions if they worked through the earlier questions quickly, whilst others were able to spend longer getting to grips with the earlier questions. Such activities have been used to ensure that all students are challenged and have resulted in high levels of engagement and motivation. Other adaptations have included the addition of thumbs down to indicate that students are unsure where to start. For example, figure 5, which displays an activity used early in the module, does not include this option whereas figure 2, an activity used later, does. Section

4.1 outlines a specific example in which the thumbs down option included in figure 2 led to scaffolding and adaptive teaching, hence this has been a valuable addition.

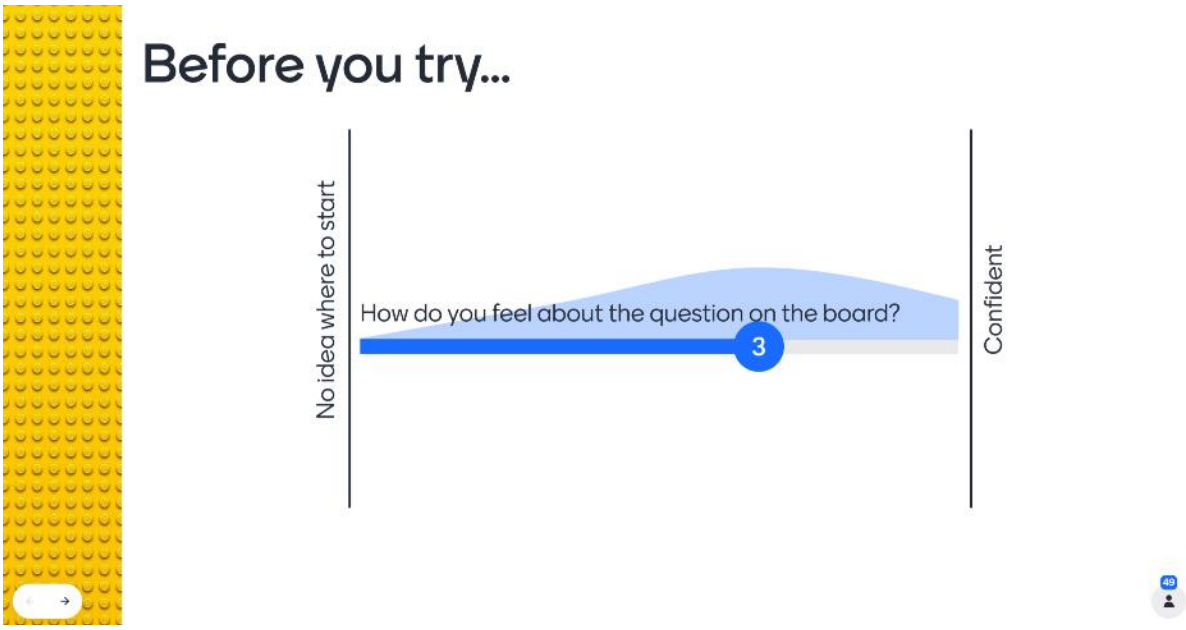


Figure 4. The first slide of an activity combining a range of strategies described in this section, specifically to gauge student confidence before attempting a similar problem, with options 1, 2, 3 and 4 corresponding to statements ‘no idea where to start’, ‘not very confident’, ‘fairly confident’ and ‘confident’, respectively.

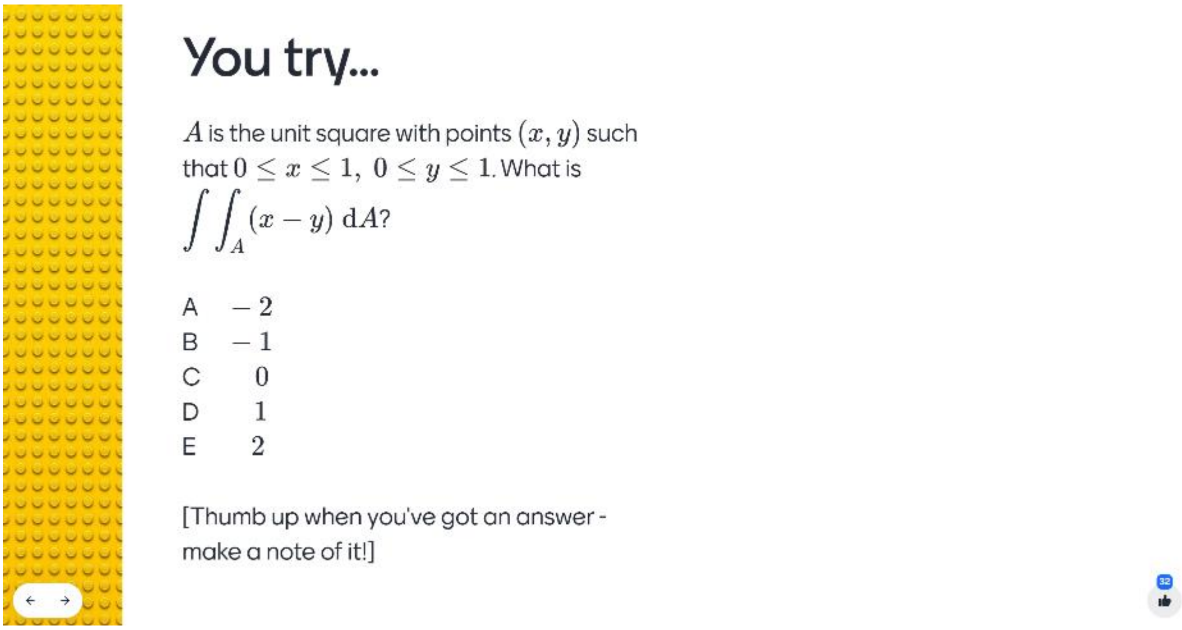


Figure 5. The second slide of an activity combining a range of strategies described in this section, specifically to gauge student progress whilst attempting a similar problem.

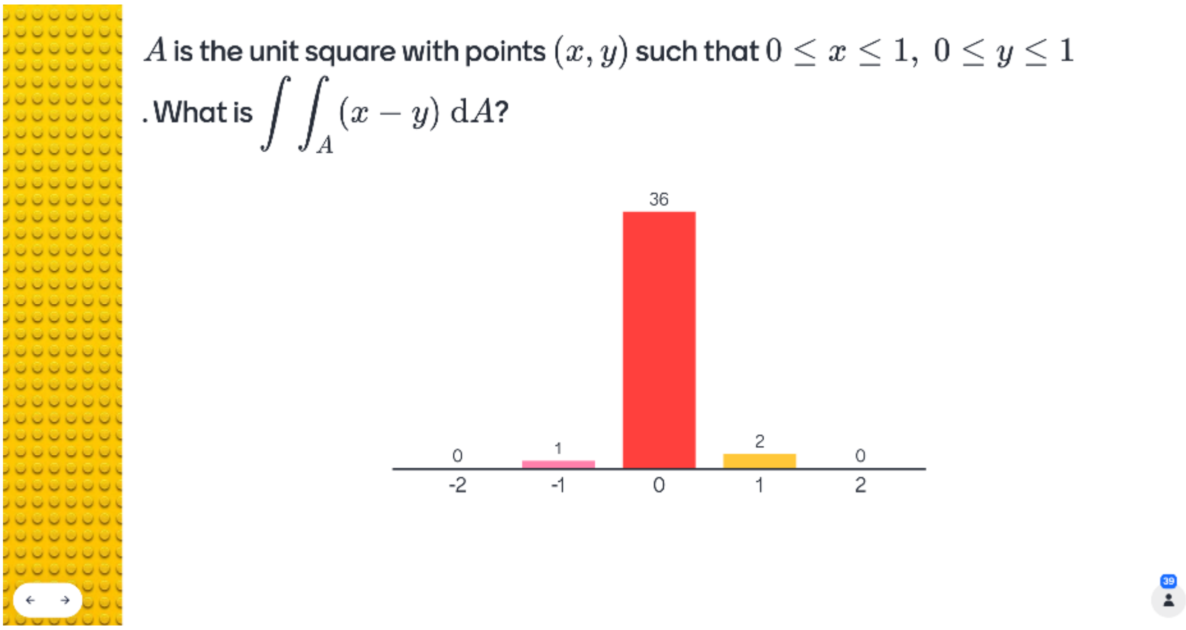


Figure 6. The third slide of an activity combining a range of strategies described in this section, specifically to gauge student understanding after attempting a similar problem.

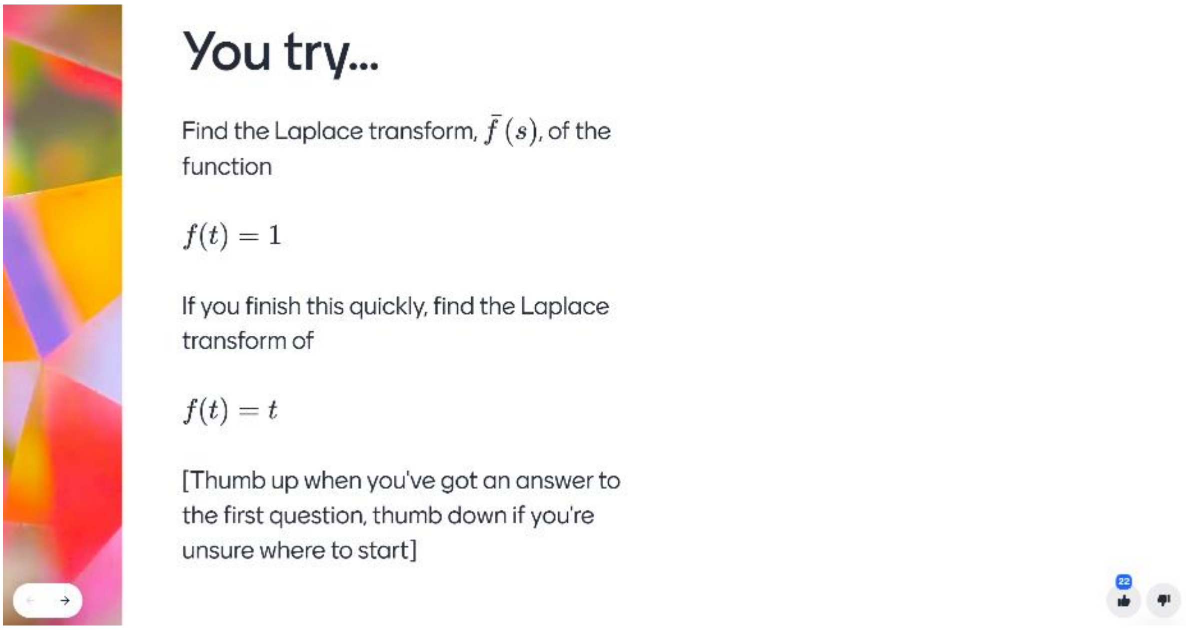


Figure 7. An example of an activity incorporating stretch and challenge via the inclusion of a second, more difficult, problem for those who completed the first problem quickly.

The activities have, in some cases, taken up more lecture time than anticipated and have thus required flexibility from the author. For the module in question, this time has been gained by removing a handful of examples from lectures and instead making them available to students on Moodle. This has worked well in the author's context. Alternatively, a flipped classroom model can be utilised, as described by Russell (2023), for example.

Finally, the activities have not been used to replace real world interactions, but as a tool to promote student engagement and to provide opportunities for formative assessment and adaptive teaching. In the author's experience, they have helped to build trust amongst students and lecturer and have encouraged real world interactions.

5. Conclusion

From the author's perspective, the interactive lecture activities have been successful in engaging students on the module; students have been actively involved in lectures and motivation has been high. The interactive lecture activities have clearly promoted formative assessment and adaptive teaching, and seemingly encouraged student ownership. The activities have been useful for time management and judging the pace of lectures also. Attendance on the module was reasonable, ostensibly better than the previous year, only dropping significantly towards the end of term.

From the students' perspective, the activities detailed in section 4 have generally been well received: specific student feedback is included below.

- "*I really like how interactive the lectures are! This is how all should be in the XXI century*"
- "*I like the interactive aspect, helps to understand the material better*"
- "*I love the interactive lecture :)*"

The module also scored highly in a 'Student Evaluation of Module' survey. The remainder of this article compares the use of activities in the module discussed thus far and a similar engineering module and outlines the author's future plans.

Student participation has been reasonable for the module discussed in this article: the author estimates that at least half of the students in attendance interacted with the activities, often considerably more, but acknowledges that this may not give an accurate representation of student understanding/confidence overall. To the best of the author's knowledge, however, students have reacted honestly, often making use of 'thumbs down' reactions to receive scaffolding. The author believes that the anonymous responses have encouraged this.

However, in a similar level 2 engineering module taught in the following semester, student participation in equivalent activities was much lower and, unlike the module discussed in this article, student buy-in was also low. The author reflects that this may be due to shared teaching on the second module: the lecture activities were introduced to students in the Spring semester after they had bought into another lecturer's teaching methods in the Autumn semester. The author persevered using the activities and highlighted their purpose when participation and engagement were particularly low, after which they did improve, but not to the extent of the original module.

Moving forward, the author aims to develop further strategies to boost student buy-in and incentivise students to engage in activities from the outset. For example, the author plans to explain their purpose at the start of a module and make explicit links to formative assessment and adaptive teaching. The author hopes that this explanation might empower students to engage in activities as a means to take control of their own learning.

In future, the author also aims to develop the activities outlined in this article to facilitate open ended questioning as another mode of formative assessment. Furthermore, the author hopes to adapt other common, engaging classroom activities for use in HE: in particular, collaborative learning strategies such as think-pair-share, group problem solving, and small group discussions. On top of the obstacles outlined in section 3, i.e. student numbers and room sizes, the layout of a standard lecture theatre is also not conducive to collaborative learning. The author hopes to develop strategies to overcome these obstacles, drawing on existing literature and their own teaching experience to formulate activities that successfully promote collaborative learning in their setting. The author is also keen to explore whether activities such as those presented in this article, inspired as they are by classroom practice, might be useful for students transitioning from school to university.

In summary, the author has aspired to re-ignite students' enthusiasm for learning in person and encourage students back to the lecture theatre in a post-pandemic world. The key strategy has been to engage students via interactive lecture activities inspired by common classroom practice, with a particular focus on formative assessment, adaptive teaching, and student ownership. These activities have created regular opportunities for formative assessment to support day-to-day learning and have promoted adaptive teaching based on students' responses: for example, prompting the author to provide scaffolding or to alter the pace of lectures to support students' needs. These features seem to have encouraged student ownership and students have taken a more active role in their learning. Student attendance has also been ostensibly better than the previous year. Overall, the author has been encouraged by the success of the interactive lecture activities introduced in this article and has frequently observed high levels of engagement and motivation amongst students.

6. References

Atkinson, R.K., Derry, S.J., Renkl, A. and Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of educational research*, 70(2), pp.181-214. https://doi.org/10.3102/00346543070002181

Black, P. and Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 92(1), pp.81-90. https://doi.org/10.1177/003172171009200119

Bloom, B. S. (1956). *Taxonomy of Educational Objectives Handbook 1*. New York: Longman.

Bloom, B.S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational researcher*, 13(6), pp.4-16. https://doi.org/10.3102/0013189X013006004

Brookhart, S., Moss, C. and Long, B. (2009). Promoting student ownership of learning through high-impact formative assessment practices. *Journal of MultiDisciplinary Evaluation*, 6(12), pp.52-67. https://doi.org/10.56645/jmde.v6i12.234

Chan, P.E., Graham-Day, K.J., Ressa, V.A., Peters, M.T. and Konrad, M. (2014). Beyond involvement: Promoting student ownership of learning in classrooms. *Intervention in School and Clinic*, *50*(2), pp.105-113. https://doi.org/10.1177/1053451214536039

Darling-Hammond, L. and Bransford, J., eds. (2007). *Preparing teachers for a changing world: What teachers should learn and be able to do*. John Wiley & Sons.

Duffy, G.G. (2006). Developing metacognitive teachers: Visioning and the expert's changing role in teacher education and professional development. In *Metacognition in literacy learning* (pp.321-336). Routledge.

Hardy, I., Decristan, J. and Klieme, E. (2019). Adaptive teaching in research on learning and instruction. *Journal for Educational Research Online*, 11(2), pp.169-191. https://doi.org/10.1080/00131911.2020.1722065

Hattie, J. (2008). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. *International Review of Education*, 57, pp.219-221. https://doi.org/10.4324/9780203887332

Mannison, M., Patton, W. and Lemon, G. (1994). Interactive Teaching Goes to Uni: Keeping Students Awake and Learning Alive. *Higher Education Research & Development*, 13(1), pp.35-48. https://doi.org/10.1080/0729436940130104

Robson, L., Gardner, B. and Dommett, E.J. (2022). The post-pandemic lecture: Views from academic staff across the UK. *Education Sciences*, 12(2), Article 123. https://doi.org/10.3390/educsci12020123

Rodger, S.H. (1995). An interactive lecture approach to teaching computer science. *ACM SIGCSE Bulletin*, 27(1), pp.278-282. https://doi.org/10.1145/199691.199820

Russell, E. (2022). Remote active learning. *MSOR Connections*, 20(1), pp.46-55. https://doi.org/10.21100/msor.v20i1.1308

Russell, E. (2023). Adapting successful online activities for in-person classes - a new challenge. *MSOR Connections*, 21(1), pp.65-75. http://doi.org/10.21100/msor.v21i1.1398

Sazhin, S.S. (1998). Teaching mathematics to engineering students. *International Journal of Engineering Education*, 14(2), pp.145-152.

Stiggins, R.J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), pp.758-765. https://doi.org/10.1177/003172170208301010

Stiggins, R. and DuFour, R. (2009). Maximizing the power of formative assessments. *Phi Delta Kappan*, 90(9), pp.640-644. https://doi.org/10.1177/003172170909000907

Swan, M. (2005). *Improving learning in mathematics: challenges and strategies*. Department for Education and Skills Standards Unit.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, 12(2), pp.257-285. https://doi.org/10.1016/0364-0213(88)90023-7

Sweller, J. and Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and instruction*, 2(1), pp.59-89. https://doi.org/10.1207/s1532690xci0201_3

Trafton, J.G. and Reiser, B.J. (1993). Studying examples and solving problems: Contributions to skill acquisition. In *Proceedings of the 15th conference of the Cognitive Science Society*, pp.1017-1022.

Vaughn, M. and Parsons, S.A. (2013). Adaptive teachers as innovators: Instructional adaptations opening spaces for enhanced literacy learning. *Language Arts*, 91(2), pp.81-93.

Wood, D., Bruner, J.S. and Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), pp.89-100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

CASE STUDY

Facilitating collaborations between home and international students; different perspectives of engagement associated with having responsibilities in a group environment

Dr Simon A. Fairfax, University of Liverpool, Department of Mathematical Sciences.
Email: simon.fairfax@liverpool.ac.uk.
Dr Subhrajyoti Saha, University of Liverpool, Department of Mathematical Sciences.
Email: subhrajyoti.saha@liverpool.ac.uk

Abstract

This case study reports on students' responses to a newly designed group project used in a second-year undergraduate module at the University of Liverpool. The goals of the group project were to improve student engagement in the module and facilitate a collaboration between home and international students to develop a broader sense of student community within the department of mathematics. The extent to which these goals were achieved are explored in detail.

Keywords: student community, student engagement, belonging, internationalisation.

1. Introduction

Higher Education in the UK has been an attractive prospect for international students, particular from China, for a number of years. In the UK sector, the most popular route to transnational education is by far joint 2+2 programmes with established overseas partnership universities. These programmes allow international students to complete two-years in their home country followed by two-years overseas, in the overseas country of their choice, where they will receive their bachelor's degree. The 2+2 programme structure creates challenges at UK institutions, such as the potential for a disconnected student community due to the lack of previous interactions among new arrivals and current students, their different academic backgrounds, and low confidence especially with language ability in a new environment. This is also an opportunity for culturally rich collaborations whereby students from varying backgrounds work together to solve mathematical problems requiring not only technical skills, but also professional skills which are highly sought after by employers. The role of module leaders, acting as a facilitator for engagement, is of paramount importance for the bringing together of the wider student community within mathematics. It is well-established that having a strong sense of student community is associated to students' academic success (Johnson et al., 2023). A student community encompasses various elements such as belongingness, connectedness, and engagement (Rovai, 2002). In 2022-23, the authors led a second-year financial mathematics module, optional on the G100 (BSc Mathematics) programme, which had 150-students. This comprised of half progressing from Year 1 and half who newly joined the University of Liverpool in Year 2 as part of a joint programme with the overseas Chinese partnership university XJTLU. Even though the University of Liverpool - XJTLU partnership was formed in 2006, this was the first year in which a group project was used in the UK to connect home and overseas students. Naturally, many of the students in each of these groups, with very different educational experiences, will have established social groups in the previous year. Without extra-curricular activities, a lack of opportunity for collaboration between home and international students within the mathematics curriculum may elicit a lack of a sense of belongingness in their community. Both groups of students have informally spoken about the disconnected nature of the cohort. The authors introduced the newly developed group project for the purpose of addressing this inherent issue.

2. Assessment design, support, and evaluation methodology

2.1 The group project assessment design

The group environment is an opportunity for students to pool experience and expertise to solve the group tasks in a more effective manner than a collection of individual contributions. After studying core modules in Year 1, home students will be on different pathway, affecting which modules they study later, varying from statistics and probability, financial mathematics, pure mathematics, and applied mathematics. XJTLU students will have studied similar modules but have a different learning experience with different educational priorities. Several academic themes were incorporated into the group project assessment design to justify the group nature of the activity.

Incorporating IT skills

One of the roles of a Financial Analyst is to value financial products which are sold by institutions in the financial services sectors; a simple example being insurance packages. Often, valuing financial products are complex so simulations requiring mathematical software are used; MATLAB is seen as the most appropriate in financial mathematics. Digital fluency is at the heart of the University of Liverpool Curriculum Framework. MATLAB and Maple are fully embedded into core Year 1 mathematics modules, in addition R training is provided to all students on the statistics and probability pathway. On the other hand, many of the XJTLU students haven't encountered MATLAB or Maple but have encountered general purpose programming languages such as Java and C++. Overall, students' expertise using mathematical software are highly varied. For one of the tasks in the group project, students were given a model for the future daily movements of a stock S of a listed company which included a normally distributed random variable R, for the daily movement, reflecting the economic backdrop of their operations. The model was, $S(t+1) = S(t)\exp(R)$, $t = 0,1,\ldots,T$, where T is the number of trading days in one-year. Students were required to re-interpret yearly parameters to obtain a value for R and run predictions for the closing share price in one year using this model. These values were then applied to a financial product whose payoff L was dependent on S. Similarly, in a different independent task, students were required to import and analyse real daily share prices for two listed companies. Once the data was stored in MATLAB, students were able to plot the daily prices, compute rolling daily rates of return and measure volitatility, investigate seasonal trends, and compute other measures of risk on return studied in the course content.

Financial mathematics – connecting content to real-life

Given the vast technical language used, students find it difficult to interpret the features of financial products. By writing an assessment task for valuing a financial product from the perspective of a fictional employer, students were presented with a different way of thinking about risk, reward, and market valuation. This new context allowed students to further explore advanced content which are used in the financial sector. A second financial product was introduced whose value is related to the original financial product via a fundamental price parity theorem in this discipline; it was up to the groups to identify relevant results from lectures. To add depth to this problem one of the essential parameters, the financial return of a notional zero-risk investment asset, was removed. This allowed students to investigate how this is interpreted in real-life, i.e., government bonds are viewed as risk-free in this topic, and present their conclusions in a formal setting.

Statistical skills – using a wide range of analytical tools

Given N predictions of the share price discussed above, students were able to take these values and calculate an average price of a complex financial product set out in the assessment. These values, however, come with a degree of uncertainty due to the sampling within the simulation. Those from a statistics background will be well-versed in, say, confidence intervals which would add depth to their analysis. Similarly, when presented with investment decisions, the coefficient of variation translates to risk-per-unit-of-return and useful for justification when making decision in portfolio management.

Research skills – research connected tasks

The parameters for R in the financial model discussed above are presented with little context. In the field of financial mathematics, one of the market models for predicting the return of an investment asset is based on an argument related to systematic and non-systematic risk, with a collection of strong assumptions. In the group project, students were required to select two listed companies, one from the UK market the other overseas, and identify and compare different types of risk from their review of the financial statements: a task with a hidden agenda of encouraging discussion between home and international students. As well as understanding these risks, students were required to present them in concise language in the report aimed at a general audience. This task had the benefit of allowing students to analyse financial institutions they find interesting and the environment in which they operate. Later, students may apply to graduate programmes for these companies hence benefit from their investigation.

2.2 Structured support

Several steps were taken by the teaching staff, authors of this case study, to support students through the potential challenges of working in a group environment; these are discussed further in this subsection.

Encouraging participation

There are different approaches of group formation discussed in the literature. Mahenthiran et al. (2000) investigated whether the performance and attitudes of students could be improved by either choosing the group randomly or combining paired friends. They observed that both student satisfaction and academic performance were significantly enhanced when students were allowed to choose a single friend in the group. This finding emphasises the importance of group formation when introducing group work. In the context of the group project, groups of size six with a combination of both home and international students in a single group was implemented. This was a modified approach to group formation, which was originally conceptualised by Mahenthiran et al. Students were given the chance to form subgroups containing three members which, in all cases, were based on social groups with common nationalities. These were then combined by the Module Coordinators into groups of six, split equally based on home and international students. This approach to group formation is based on Bradshaw (2009). Groups were formed one week before the start of the assessment activity and given organisational tasks to allow them to introduce themselves to each other before the assessment activity as well as become familiar with the online environment. One of the draw backs in this group allocation was it had little influence relating to group dynamics, e.g., ensuring leaders and other relevant characteristics are present. Every group was required to appoint a group leader and while each group was successful in identifying a leader, it is important to acknowledge the potential for a lack of willingness among group members to assume this leadership role. In order to mitigate this issue, one may wish to implement a mechanism that ensures the presence of leadership and other relevant characteristics within each group. In this context, the study conducted by Rowlett (2013) offers some valuable insights.

Supporting general communication

To support general communication within groups, private channels were set-up on Microsoft Teams allowing students to conveniently connect via message and video call as well as share documents.

Each group was required to appoint a Group Leader to ensure the group remained on task and to report any issues outside of their control, e.g., report weekly absences to the Module Coordinator.

Assisting with direction

Groups were required to hold at least one group meeting every week with all members present. As part of these meetings groups were required to agree upon plans to divide-up tasks, establish subgroups based on preferences, strengths, and experience, agree provisional deadlines to tasks etc. A requirement of the group project was for students to submit the minutes of their weekly meetings.

Promoting engagement

Employer themed projects whereby students are required to address their findings to a mock employer in a mock secondment scenario are an effective means to promote engagement and raise confidence (see Fairfax, 2023). For many students, this will be their first encounter with leadership and group accountability as part of their studies.

Recognising unequal contributions:

As recently highlighted by Shaw (2022), one common gripe with group work is the perceived lack of fairness when group members do not contribute equally. For simplicity, a group can receive one collective grade and each member receives the same grade. This is often fitting since, in most cases, every group member contributes similarly to the project's outcome. Yet, there are rare instances where one person might not contribute as much as the others. To address this issue, which could lead to wider dissatisfaction, clear instructions indicating individual and group responsibilities were written and regularly sign-posted during the assessment activity. At the conclusion of the group project, each student was required to complete a peer moderation task. A popular mechanism of peer moderation was identified in the project EPCoS based on practice at the University of Exeter (Milne, 1998). This approach was followed in the group project but conducted digitally. Each student had to complete a 'Buddy Check' (licenced software that can be integrated with learning management system) appraisal questionnaire, ranking the contributions of each group member as part of the peer moderation. This method helped the Module Coordinators pinpoint those who have either contributed less or more than intended by the group. Any deviations from the group average were reconciled with weekly group meeting notes and individual marks were adjusted according to contributions. The goal isn't to make detailed distinctions among group members but to acknowledge notably high or low contributions, e.g., when group work doesn't go according to plan because of non-participation, by adjusting individual marks. A similar method for adjusting the marks following peer moderation is discussed by Ian (1999).

2.3 Method for evaluation

A modern model of student engagement presented by Borup et al. (2020) sets out three indicators of engagement: affective (or emotional), behavioural and cognitive engagement. Emotional engagement is associated with emotional involvement in learning such as enjoyment, confidence and satisfaction or at the opposite end, boredom, frustration and dissatisfaction. Whereas behavioural engagement involves physically completing tasks associated with learning with, say, attending and participating in activities viewed as low and high levels of engagement respectively. In this study, several potential barriers were identified by the authors and students were asked to reflect on:

- belongingness at university,
- language confidence,
- satisfaction with their studies,
- their perceived academic readiness relative to their international peers, and

- whether academic ability was a factor in communication.

Cognitive engagement involves productive involvement with learning activities and viewed as the higher level of engagement associated with questioning and exploring concepts. The authors identified the following key areas for investigation: effort put into studies, self-regulation of study, direction with study, and perceived levels of engagement with studies. At the conclusion of the project, all students were invited to complete a questionnaire associated with engagement which are discussed in detail in section 3.

3. Students' responses and discussion

3.1 Goal 1: Facilitate collaboration between home and international students

There will be range of different emotions when a group project is introduced to students from diverse cultural backgrounds. Some students will want to connect with their international peers, but not have the confidence or feel that they have had an opportunity to make an approach to linking up. Other students may be less open to collaborating due to awkwardness or shyness, say, potentially due to feeling unprepared because of language confidence or perceived academic ability, hence prefer individual assignments and the comfort of their social groups. In this group project, only 3 out of the 25 groups appointed a Project Leader who was an international student highlighting the challenges associated with confidence among these students. In terms of backgrounds, the home students have encountered numerous group projects at the University of Liverpool within their first year of study, on the other hand, international students have been exclusively assessed using time constrained examinations (with some lower stakes in-semester continuous assessment) at XJTLU in their home country presenting challenges with group dynamic. It is important to also recognise different priorities within the student community. For example, compare students: (a) who are preparing for a career in finance or similar, with an international corporation in mind, who seek opportunities to work on a real-life project with new people; and (b) who intend to go onto postgraduate studies who will be looking to take on individual research projects and maximise their overall module mark to improve their applications. Those in the latter category may feel group work, particularly with less motivated group members, is unfair and prefer closed-book examinations. On the other hand, students from the former category will likely prefer the group set-up because they see its value, for example to demonstrate professional skills, as they seek to break into the graduate job market. Regardless, an ability to work with peers and communicate mathematics, orally or as a part of formal report writing, are essential skills for all students. The group project was designed in a way to allow students to establish their own workflows, creating personal accountability and requiring active involvement for all members. In reality, based on open responses about their positive experiences during the project as well as viewing subgroup formations, collaboration fell into two categories.

Category 1: Active collaboration with international peers involving problem solving, the widening of social groups and raising cultural awareness.

At the planning stages, a total of 10 out of the 25 groups, or 40% of students, discussed their academic experience prior to the project, identifying their strengths, weaknesses, and areas they wish to develop, and accordingly established sub-groups that best utilised the overall group's expertise. This contrasted with the remaining students who, on face-value, divided tasks by social groups, i.e., their nationalities, where communication skills and language ability were barriers. Those groups who actively collaborated with international peers wrote about their useful experiences in the open response section.

> *"Although awkward at first studying and problem solving with foreign peers was an eye opener and I believe ultimately useful in the long run, odds on at some point in my future professional career I will have to work along side someone who's first language is not english, or even I'm*

in a foreign country where I am the minority. Hopefully even this small taster of that will aid me in the future."

"Good to communicate and work together with international students, gained experience which we can talk about in the future"

"It was interesting working with people who have clearly been taught different previously, some of their knowledge was amazing."

"Collaboration and the availability of different minds when facing challenges"

"It was fun to meet and get to know more international students and work with them on the project."

These comments highlight the value attributed to facilitating collaborations between home and international students. These students were open to the challenges and persisted showing adaptability during tough times. There is evidence of peer learning on an academic level accompanying the professional skills development in terms of international awareness, as well as student satisfaction.

Category 2: Passive collaboration with international peers and raising cultural awareness

A total of 15 out of the 25 groups, or 60% of students, decided from the outset that they would work in subgroups which were based on their nationalities rather than, it seems, to work to their academic strengths. Many of these groups explicitly mentioned in their meetings notes that language was a challenge with group communication. These groups attempted to address these challenges with native speakers adjusting the pace of their speech and non-native speakers seeking clarification where there were misunderstandings; unfortunately, this did not transpire into active problem solving between international peers. Students' comments about language as a barrier are discussed in section 3.2. Even though these subgroups were formed based on their nationalities, at the weekly team meetings all students were required to come together and discuss their progress. There is some evidence of wider home-international student collaboration at this stage, for example, one student wrote:

"Personally I was assisted and assisted one of the Chinese students so we interacted a fair but, however I am not sure the same can be said for the rest of the group."

This situation demonstrates individual willingness for wider collaboration. To address the passive collaboration issue between home and international students in future years, teaching coordinators have identified student volunteers to share their experience of the project who will provide written advice for future students based on their own experiences in the current year. One way for the facilitators to encourage the whole group to think more careful about their individual skills when establishing sub-groups is to consider the number of independent tasks in the group project. By having a project which naturally splits into two independent parts has translated into this 2 : 3 active-passive collaborative split. It remains to be seen in future years whether having, say, three independent tasks within the project, with a strong student voice from the previous year, will coax more students into the active category.

3.2 Goal 2: To improve engagement particularly between home and overseas students

In section 2.1, five perceived categories influencing students' emotional and behavioural engagement were set out. Similarly, for the sub-components of cognitive engagement. As part of the evaluation, students were asked to reflect upon changes within each of these categories after the group project had run and before receiving their grades and feedback; a detailed analytical review will follow at a later date. In the open response section, students were invited to write about any positive experiences; these are explored in the remainder of this case study.

Emotional and behavioural engagement

Generally speaking, students who feel disconnected from the student community are less likely to be engaged with their studies and hence less likely to play a meaningful role in group work. Students' sense of belonging at an institution shouldn't be taken for granted. Prior to the group project, both home and international students expressed feelings of disconnectedness; this being a major barrier for student engagement on an emotional level. When asked to comment on positive experiences from the project, two responses related to belonging.

> *"Meeting students i wouldn't necessarily have spoken to before."*

> *"It was fun to meet and get to know more international students and work with them on the project."*

Even though the group project had influence in this direction, there is more to be done to revive the wider sense of student community in the post lockdown period, for example, through student societies at the university level.

Language ability will naturally be a challenge for all non-native English speakers and this group project was no exception. Overseas students find themselves in a dominant English-speaking environment and may refrain from discussions due to being misunderstood or not being able to keep up with native speakers. Secondly, native speakers need to be aware of varying language abilities, otherwise they may be less willing to discuss tasks if they feel group members don't understand them. Several groups specifically mentioned language as a barrier to collaboration in their weekly meeting notes which goes some way towards accounting for the lower than expected active-passive collaborative split discussed in section 3.1. The group environment was an opportunity for all international students to put their language skills to use as well as a self-realisation exercise to discover any shortcomings in their abilities. For home students, it was also an opportunity to encounter group dynamic whilst working in diverse teams. Several comments were made in the questionnaire on this theme.

> *"Learning to communicate with international students, especially realising our grammar and word choices are different."*

> *"I have learnt what it is like for international students studying at this university in a new country and the language challenges they must face."*

> *"Working with international peers, I have a better understanding of their speech and habits."*

> *"Shows us how the international students experience university and the way they tackled the group project."*

> *"Having to deal with and overcome language barriers."*

Unsurprisingly, the low level of language confidence was skewed by non-native English speakers. Practice makes perfect and more opportunities to put language skills, and other broader communication skills, to the test facilitates this practice. The above comments suggest the group project has set an important foundation for bringing the wider student community together in this context.

Cognitive engagement

The project required students to use and develop diverse skill set and apply them appropriately in solving the open natured tasks. This includes but not limited to IT skills like programming, knowledge of financial mathematics and statistics, etc. As a result, students need to put in effort to not only understand their own areas of expertise, but also to grasp the basics of other areas to effectively engage when problem solving with their peers. Several comments were made in the questionnaire about positive experiences on these themes.

> *"Using Matlab to refresh my memory"*

> *"It is necessary to consider the problem from the other side's point of view"*

> *"As for the code, it is necessary to provide corresponding comments and explanations, rather than directly submit the finished part. It is necessary to consider the problem from the other side's point of view"*

> *"Collaboration and the availability of different minds when facing challenges"*

> *"Meeting new people. Working together"*

As discussed earlier this project was aimed to mimicking a real-life scenario. Consequently, this has allowed students to see real-life connections and hence provide direction as they progress with their studies. Several comments were made in the questionnaire on this theme.

> *"Applying coding to real world scenarios."*

> *"Specialisation of tasks allowing everyone to work to their strengths"*

> *"Having the group project relate to potentially an actual real world problem was very useful in terms of putting things into perspective of how approaching and completing a problem may work in the real world"*

4. Conclusion

This case study aims to demonstrate the potential of structured group work. The group project's design, emphasising real-world applications and the use of mathematical software were not only aimed to enhance the mathematical/technical proficiency of the students but also aimed to foster a sense of belonging and community; a major challenge given the nature of 2+2 programmes. As discussed above, the structured support provided by the teaching staff played a crucial role in facilitating effective communication and collaboration among students. The study revealed a range of emotional responses to this initiative. The feedback from students' post-project was interesting; a small number of groups openly expressed a newfound sense of belonging and appreciation for the diverse perspectives brought by their international peers. While some international students actively sought to connect with their home peers, others were hesitant, often due to language barriers or a lack of confidence in their academic abilities. Several of these students expressed regret not communicating with home students more as they prepare for English Language proficient tests required for postgraduate studies, but struggle to practice oral English. A small group of these students have agreed to provide advice for future students, from the experienced student voice, to help them overcome perceived barriers to engagement in future years. The insights gained from this study may be valuable to other educators and institutions seeking to create more inclusive and collaborative learning environments and both authors would welcome opportunities for their input.

5. References

Beer, C., Clark, K., and Jones, D. (2010). Indicators of engagement. *Proceedings ascilite Sydney*, pp.75-85. Available at: https://ascilite.org/conferences/sydney10/procs/Beer-full.pdf [Accessed 23 April 2024].

Borup, J., Graham, C. R., West, R. E., Archambault, L., Spring, K. J. (2020). Academic Communities of Engagement: an expansive lens for examining support structures in blended online learning. *Educational Technology Research and Development*, 68, pp.807-832. https://doi.org/10.1007/s11423-020-09744-x

Bradshaw, N. (2009). Tribute to triplets: A model for successful group work. *Compass: The Journal of Learning and Teaching at the University of Greenwich*, 1, pp.7-9. Available at: https://journals.gre.ac.uk/index.php/compass/article/view/9/21 [Accessed 23 April 2024].

Fairfax, S. A. (2021). Improving student engagement through employability themed group work. *MSOR Connections*, 21(1), pp.23-28. https://doi.org/10.21100/msor.v21i1.1396

Filade, B. A., Bello, A. A., Uwaoma, C. O., Anwanane, B. B., and Nwangburuka, K. (2019). Peer group influence on academic performance of undergraduate students in Babcock University, Ogun State. *African Educational Research Journal*, 7(2), pp.81-87. https://doi.org/10.30918/AERJ.72.19.010

Greene, B. A. (2015). Measuring Cognitive Engagement With Self-Report Scales: Reflections From Over 20 Years of Research. *Educational Psychologist*, 50(1), pp.14-30. https://doi.org/10.1080/00461520.2014.989230

Johnson, M.D., Margell, S.T., Goldenberg, K., Palomera, R. and Sprowles, A.E. (2023). Impact of a First-Year Place-Based Learning Community on STEM Students' Academic Achievement in their Second, Third, and Fourth Years. *Innovative Higher Education*, 48, pp.169-195. https://doi.org/10.1007/s10755-022-09616-7

Mahenthiran, S. and Rouse, P. (2000). The Impact of Group Selection on Student Performance and Satisfaction. *International Journal of Educational Management*, 14(6), pp.255-65. https://doi.org/10.1108/09513540010348043

Rovai, A.P. (2002). Building Sense of Community at a Distance. *The International Review of Research in Open and Distributed Learning*, 3(1). https://doi.org/10.19173/irrodl.v3i1.79

Rowlett, P. (2013). A Modification of Bradshaw's Method of Group Allocation When You Do Not Know the Students. *MSOR Connections*, 13(2), pp.43-50. Available at: https://web.archive.org/web/20210329124904id_/https://s3.eu-west-2.amazonaws.com/assets.creode.advancehe-document-manager/documents/hea/private/msor_13-2-7_1568037398.pdf [Accessed 23 April 2024].

Shan, Li. (2021). Measuring Cognitive Engagement: An Overview of Measurement Instruments and Techniques. *International Journal of Psychology and Educational Studies,* 8(3), pp.63-76. Available at: https://dergipark.org.tr/en/pub/pes/issue/64305/976270 [Accessed 23 April 2023].

Shaw, L. M. (2021). "It's so unfair" – Can we increase student perceptions of equity in the grading of group assessments by allowing them to declare a distribution of workload? *MSOR Connections* 21(1), pp.23-28. https://doi.org/10.21100/msor.v21i1.1376

Trowler, P. and Trowler, V. (2010). *Student Engagement Evidence Summary*. The Higher Education Academy. Available at: https://www.advance-he.ac.uk/knowledge-hub/student-engagement-evidence-summary [Accessed 23 April 2023].

Utting, I. (1999). Negotiated assessment criteria and peer assessment in software engineering group project work: A case study. In J. Daudt and O. Rompelman (Eds.), *What have they learned? Assessment of Student Learning in Higher Education*, pp.165-170. Brussels: SEFI. Available at: https://kar.kent.ac.uk/21848/ [Accessed 23 April 2023].

Milne, W. (1998). *Moderation Using Student Input*. Project EPCoS internal document. Available at: https://www.cs.kent.ac.uk/national/EPCOS/ [Accessed 23 April 2023].

CASE STUDY

Toward a holistic approach to mathematics support at Munster Technological University's Academic Learning Centre

Marian Hurley, Department of Marketing and International Business, Munster Technological University, Cork, Ireland. Email: Marian.Hurley@mtu.ie
Deirdre Casey, Academic Learning Centre, Munster Technological University, Cork, Ireland. Email: MarieD.Casey@mtu.ie

Abstract

This case study outlines plans for a revised approach to mathematics support at Munster Technological University's Academic Learning Centre. Under this proposed approach, mathematics support – delivered both on a one-to-one basis and in group settings – is to be supplemented by and integrated with learning supports offered by the Academic Learning Centre's sister programmes, Academic Success Coaching and Navigate Learning Development. This new approach will see the three programmes brought together under the Academic Learning Centre name. In implementing this approach, we have two key objectives: firstly, to determine the impact of this integrated approach on our mathematics support service and on its users – particularly those displaying traits of mathematics anxiety – and, secondly, to assess the extent to which presenting students with one integrated support option may result in increased service access and use.

Keywords: mathematics support, academic coaching, approaches to mathematics support.

1. Introduction

Mathematics support has operated as part of the Academic Learning Centre (ALC) in Munster Technological University (MTU) since 2005. In more recent years, two additional student success programmes have been established at MTU. These are Academic Success Coaching (ASC) and Navigate Learning Development (Navigate). All three are programmes of AnSEO, the university's Student Engagement Office.

Since the university's post-COVID return to on-campus activity in 2021-22, staff of these three programmes have been co-located in the ALC's hub. Due to a natural increase in collaboration between the areas, it soon became clear that our objectives and methods were complimentary, and that the synergies of approach represented a particular opportunity to address the priority areas of mathematics anxiety and mathematical resilience. Consequently, in June 2023, the three programmes were merged to form a new, rebranded Academic Learning Centre.

The rationale for this development is two-fold: first, it aims to address the real and prevalent mathematics challenges faced by our students in a holistic and developmental rather than a fragmented and episodic fashion. It also aims to streamline and simplify student- and staff-facing messaging regarding academic supports available in the institution, with a view to removing barriers to students who may know they need some academic support but may not know which of the three strands represents the best starting point for them.

In this case study paper, we first discuss the academic supports which were available to students prior to June 2023. We go on to discuss Mathematics anxiety and resilience and their possible connections to coaching in higher education, before finally outlining key reasons for implementing a change and describing our new model. To conclude, we highlight some early indications of success.

2. Background

The ALC at MTU - formerly Cork Institute of Technology (CIT) – is widely known and recognised by students and staff of the university's Cork campuses. It provides individual and group supports across a range of subjects, of which mathematics and statistics have traditionally been the largest elements with the highest student demand. The ALC's services are available to all students free of charge. While its activity is co-ordinated by one dedicated staff member, a distinguishing feature of the centre is the fact that supports are delivered by MTU lecturers who are allocated a number of hours per week. This interface with lecturers is an attractive proposition for students, and allows meaningful collaboration and cross-pollination between academic departments and the ALC. As such, the ALC is not a bolt-on but a living part of programme delivery in the university.

ALC supports operate on an opt-in basis. Students are informed about the centre and its website (https://studentengagement.cit.ie/alc) through in-lecture workshops facilitated by ALC staff, through all-student emails and regular social media posts on three AnSEO-administered Instagram accounts, through lecturer recommendations, and, often, through peer signposting. Any student wishing to attend can book a 30-minute one-to-one (or small-group) appointment in person or online either by visiting the ALC hub and requesting a booking or, more commonly, by using the booking calendar available through the ALC website. Drop-in support sessions which can be attended by multiple learners are also available.

In semester 1 of 2022-23, the ALC offered 16 hours of mathematics support per week across one-to-one appointments, small-group appointments and drop-in hours. An ALC attendee will typically bring a topic, question, or task by which they are challenged to their session, and they will be facilitated to address it with the lecturer. Follow-up appointments are encouraged, and return visits are common. For example, in semester 1 of 2022-23, 144 students attended for 458 mathematics support visits. Over the lifetime of the centre, mathematics lecturers have noted that many students have appeared to exhibit signs of mathematics anxiety (on which more follows). More specifically, this has been a recurring theme for students of non-technical and/or non-STEM degrees whose progression from one stage to the next depends on passing a mathematics module.

ASC has traditionally provided bookable one-to-one coaching sessions or in-lecture group coaching workshops. Similar to ALC supports, students can opt in to an online or an in-person one-to-one coaching session by visiting the ALC office and requesting an appointment or by booking through the booking calendar accessed through the coaching website (https://studentengagement.cit.ie/asc). The programme is overseen and delivered by two full-time coaches. ASC workshops are usually requested by lecturing staff to address a theme or area of particular relevance to a student group. Both one-to-one sessions and group interventions typically target themes like motivation, goal-identification and goal-setting, and identification and exploitation of personal strengths. Student coachees generally work with their coach to identify actions which will enable them to achieve a goal or goals which they have identified as important to them.

Navigate focuses on generic skills for university studies. It, too, interacts both with individual students and with whole-class groups. The programme was developed in response to requests from both lecturing staff and students for additional supports in areas such as writing and presentation, but also in disciplinary competences. In the latter area, workshops covering, for example, technical and scientific writing skills have been piloted. In 2022-23, through a combination of group and one-to-one interventions, Navigate interacted with students more than 1,450 times.

The separate branding and messaging of the programmes challenged students and staff. Furthermore, it was difficult clearly to delineate the remit of each for a growing student body with limited time to interact with additional supports. As such, our objectives for the programme merger are two-fold: they are, firstly, to determine the impact on our mathematics support service and its users of this integrated approach (with specific reference to working with mathematics anxiety) and, secondly, to assess the extent to which service access and use increases as a result of streamlined messaging. It is anticipated that communicating the availability of comprehensive support for any academic issue students may be confronting will enable learners to find our services more readily.

3. Mathematics anxiety, mathematical resilience and coaching: an overview

The emerging discourse around mathematics anxiety and mathematical resilience is at the forefront of our thinking. Mathematics anxiety has been recognised as a phenomenon impacting students' mathematical performance for some decades. As early as the 1980s, the impact of early negative experiences with mathematics on adults' mathematical self-concept and avoidance of numerical challenges was discussed in the literature of second-level teaching (Morris, 1981). However, more recent scholarship has pointed to the fundamental importance of awareness of mathematics anxiety and approaches to working successfully with mathematics-anxious learners on effective mathematics learning and teaching (and support) in higher education (Núñez-Peña, Suárez-Pellicioni and Bono, 2013; Núñez-Peña, Bono and Suárez-Pellicioni, 2015; Bhowmick et al., 2017). As such, we are curious to determine how an intentional integration of mathematics support, ASC and Navigate could impact the effectiveness of the ALC in facilitating students to confront mathematics anxiety.

Below, we provide an overview of the salient features of both coaching and of mathematical resilience within higher education. We seek to identify the areas in which coaching can support mathematical resilience, thus potentially addressing emerging priorities in higher education mathematics support.

3.1 Mathematics anxiety and mathematical resilience

Much scholarship has been devoted to the incidence of and learning challenges posed by mathematics anxiety, especially among students of disciplines not specifically connected with numeracy and among students belonging to 'non-traditional' demographic groups. The literature suggests that more mature undergraduates may experience lower mathematical self-efficacy than their younger counterparts (Jameson and Fusco, 2014; Marshall et al., 2017). Bull (2009) investigated the prevalence of mathematics anxiety among nursing students, and Chapman (2010) reported on the impact of a foundation mathematics course using an applied approach for students of geography. The impact of mathematics anxiety on business undergraduates, too, has received comment (Joyce et al., 2006), and it has been argued that a holistic approach is required to address the mathematical self-concept and self-efficacy issues that can arise for students of marketing in higher education (Bhowmick et al., 2017). Núñez-Peña and Bono (2021) have elucidated the negative impact on mathematical performance in multiple choice tests where a student exhibits both high mathematics anxiety and high levels of perfectionism. Coaching has been cited as a constructive approach to the cultivation of the self-efficacy that mathematics-anxious learners may need to develop (Lee and Johnson-Wilder, 2017).

Cousins et al. (2019) have suggested that resilience comprises two facets: first, the experience of being overstretched by challenge and second, the ability to manage the emotion which accompanies it. Drawing on Dweck's (2000) concept of the growth mindset, Lee and Johnston-Wilder (2017) have proposed the "*growth zone*" approach to tackling mathematics anxiety through a specifically "***mathematical*** *resilience*" (emphasis ours) which recognises the impact of the emotional experience arising from mathematical challenge. They theorise that this mathematical resilience comprises four facets: the growth mindset, the value placed by the learner on mathematics, an understanding of the process of and perseverance needed to learn mathematics and an understanding of how to seek support and collaborate on mathematics tasks (Lee and Johnston-Wilder, 2017). Developing mathematical resilience, they argue, enables the student to persevere in the "*growth zone*" which represents the middle ground between working without challenge on familiar tasks and becoming overwhelmed by emotional responses to mathematical challenge in a "*danger zone*" (Lee and Johnston-Wilder, 2017).

3.2 Coaching in Higher Education

The International Coaching Federation (ICF) defines coaching as "*partnering with clients in a thought-provoking and creative process that inspires them to maximize their personal and professional potential*" (ICF, 2023). Although nebulous, the definition points to the centring of the coaching client in the process and to the accent on growth and development. Per Whitmore (2009), coaching is focussed on developing people's capacity to learn rather than on teaching.

In her monograph on coaching in higher education, Andreanoff (2016) indicates that the facilitative and non-directional nature of coaching is suited to building self-efficacy. She argues that using coaching techniques minimises the need for advice to be offered to help-seeking learners, thus scaffolding them to identify their own path to academic success. Therefore, the potential for a collaboration between mathematics lecturers and coaches to develop mathematical self-efficacy is clear. Spencer (2021) notes that with such measures in place, coaches working with "*non-traditional*" students in higher education will need to be attentive to finding a suitable equilibrium between support and challenge, mirroring the "*growth zone*" approach.

Coaching is explicitly concerned with the development of resilience. Of the positive psychology approach to coaching, Boniwell and Kauffman (2018) write that its roots are in the adult learning field, and it is inflected with a variety of psychological traditions focussing on the trajectory from self-efficacy to goal-achievement. Like Lee and Johnston-Wilder (2017), Boniwell and Kauffman (2017) highlight the role of positive emotion in enhancing performance and achieving measurable outcomes. They also emphasise the power of improving outcomes further to bolster positive affect.

While recognising its potential, we do not seek to endorse the "*coaching approach*" unproblematically. Some of the concepts which have informed the discipline – notably the "*growth mindset*" itself – have come under scrutiny and remain the subject of debate. The possible misapplications of the growth mindset in adult mathematics education have been signalled by Barger, Xiong and Ferster (2022). Yan and Schuetze (2023) have highlighted the ambiguity of the concept and of its measurement and indicate that assessing its utility conclusively is challenging. Indeed, Dweck herself returned to the concept, acknowledging and responding to the controversies surrounding it in a 2020 article co-authored with David S. Yaeger (Yaeger and Dweck, 2020).

Notwithstanding valid critiques of the approach, in the context of the new ALC, the coaching approach, with its focus on resilience and self-efficacy, merits piloting as a potentially relevant means of working constructively with students on issues surrounding mathematical self-efficacy.

4. Impetus for change at the ALC

During the initial post-COVID transitional period, it quickly became apparent that a significant number of students seeking academic support at MTU had multifaceted requirements. Co-location of our support programmes made the process of cross-referral and collaboration on student cases more seamless and brought our attention to the possibilities of an integrated model. Closer collaboration between the university's core student-facing academic support programmes thus emerged as a critical priority. Their integration will, we argue, create a 'safety net' for students facing academic challenges. Furthermore, students interacting with each programme throughout the transition back to on-campus learning post-COVID tended to report broadly similar challenges in re-engaging with traditional university life which called for a complex support plan drawing on the expertise of staff from all three areas. The combination of these factors amounted to a compelling rationale for full integration of the programmes under the best recognised of the three programme titles, the ALC.

The streamlining of both service and message is also likely positively to impact the delivery of ALC supports. By addressing the previous fragmentation of academic supports, we expect to:

- reduce time spent by students differentiating between programmes and by us explaining the discrete remit of each to both staff and students;
- increase the value and effectiveness of time spent in classroom visits, allowing more time for content delivery or task facilitation;
- eliminate confusion caused for staff and students by our communication of three separate messages and, as a result, removing a significant barrier to students' engagement with us;
- eliminate 'cracks' through which a student might fall when, for example, being referred to mathematics support by a coach without a mechanism for accountability or follow-through; and
- design and implement a holistic, coaching approach to mathematics supports to provide a meaningful response to the incidence of mathematics anxiety discussed above.

5. A new model

Under the integrated model, launched in September 2023, a student presenting to the ALC will no longer need to determine whether their academic performance limitations stem from the challenges posed by an individual module or topic, from broader, underlying study skills or study management needs, or from past negative experiences of the subject. In the case of a student facing difficulties in a mathematics module, for instance, an initial approach to the ALC could result in a one-to-one mathematics support appointment being booked, but where the mathematics tutor determines that there are underlying challenges with key skills such as note-taking or with clarity around learning goals, the onward referral to Navigate or ASC can be immediate and concrete rather than a suggestion.

Collaborative work between tutors, coaches and learning developers opens the way for new possibilities in the provision of one-to-one mathematics support and, especially, in the cultivation of mathematical resilience. It will represent a move away from working episodically with students on specific concepts or problems and toward working holistically with students to get to the bottom of their mathematical challenges.

The current pilot will allow us to determine:

- whether or not the new approach increases student uptake of mathematics support;
- the extent to which students find an integrated approach to mathematics support (more) effective;
- whether by amalgamating and integrating services and streamlining messaging, students find and engage with the ALC more readily;
- the extent to which students report meaningful decreases in mathematics anxiety and increase in mathematical resilience; and
- where further scope for meaningful cross-discipline offerings may exist.

To support embedding of the pilot model, a new mode of interaction between student and ALC is currently being trialled. Through discussion of the best mechanisms through which to deliver the integrated model, it was agreed that offering a bookable, 20-minute, one-to-one initial consultation for any student wishing to avail of academic would practically support our holistic approach. This has been entitled 'Academic Mentoring', is open for bookings via the ALC's online booking system, and has attracted significant engagement from students since its introduction in September 2023 (100 students during the first three months). That this option has proven popular justifies the choice to remove the decision barrier students previously faced.

6. Discussion

By creating opportunities for staff to work together on the same students' cases, we aim to generate nuanced perspectives on the challenges and mathematical anxieties learners are reporting upon initial contact with the ALC. This will drive ongoing development by highlighting areas in which additional interventions or new intervention types are needed. We also anticipate that by looking at individual student cases and at the broader picture regarding student need and priorities, we will more quickly and more effectively identify gaps in our provision.

This process of gap identification and intervention design would be aided by a higher influx of students, which should arise from simplified messaging. Initial indications regarding student uptake are positive, with an increase of 12% on the numbers of students accessing mathematics support in semester 1 2023-24 when compared to the same period in 2022-23. Once embedded, the new model, with its single logo and single message should position the ALC as the single, easily identifiable hub for academic and mathematical support of any kind at MTU. This should position the ALC as the first port of call for any academic issue and should highlight to staff the simplicity of referring a student whom they perceive to be in need of additional support.

7. Conclusion

Ethical approval will be sought to run interviews and focus groups with learners who have interacted with the ALC throughout the pilot year. The participant sample will comprise students who have engaged with the ALC, ASC and/or Navigate pre-amalgamation, learners who have engaged only post-amalgamation, and those who have experienced both models. Sampling will reflect the demographic profile of the university's student body as closely as possible.

When evaluating impact, we will be focussed on the effects of increased mathematical resilience training, on the student experience of receiving support from more than one discipline and on the influence of the holistic approach on the student's decision to return or not to return to the ALC. We will further investigate any impact students report in relation to their attitudes to mathematics challenges and their mathematical resilience. Analysis of engagement and re-engagement trends will be used to determine the real impact of integration of support disciplines on students' academic success and outcomes.

8. References

Andreanoff, J. (2016). *Coaching and Mentoring in Higher Education: A Step-by-Step Guide to Exemplary Practice*. London: Palgrave Macmillan.

Barger, M.M., Xiong, Y. and Ferster, A.E. (2022). Identifying false growth mindsets in adults and implications for Mathematics motivation. *Contemporary Educational Psychology*, 70, pp.1-16. https://doi.org/10.1016/j.cedpsych.2022.102079

Bhowmick, S., Young, J.A., Clark, P.W. and Bhowmick, N. (2017). Marketing Students' Mathematics Performance: The Mediating Role of Math Anxiety on Math Self-Concept and Math Self-Efficacy. *Journal of Higher Education Theory and Practice*, 17(9), pp.104-117.

Boniwell, I. and Kauffman, C. (2018). The Positive Psychology approach to coaching. In E. Cox, T. Bachkirova and D. Clutterbuck (Eds.), *The Complete Handbook of Coaching*, 3rd ed. London: Sage (pp.153-166).

Bull, H. (2009). Identifying maths anxiety in student nurses and focusing remedial work. *Journal of Further and Higher Education*, 33(1), pp.71-81. https://doi.org/10.1080/03098770802638689

Chapman, L. (2010). Dealing with Maths Anxiety: How *Do* You Teach Mathematics in a Geography Department? *Journal of Geography in Higher Education*, 34(2), pp.205-213. https://doi.org/10.1080/03098260903208277

Cousins, S., Brindley, J., Baker, J. and Johnston-Wilder, S. (2019). Stories of Mathematical Resilience: How Some Adult Learners Overcame Affective Barriers. *Widening Participation and Lifelong Learning*, 21(1), pp.46-70. https://doi.org/10.5456/WPLL.21.1.46

Dweck, C. (2000). *Self-Theories: Their Role in Motivation, Personality and Development*. New York: Psychology Press.

International Coaching Federation (2023). *All things coaching*. Available at: https://coachingfederation.org/about [Accessed 11 November 2023].

Jameson, M.M. and Fusco, B.R. (2014). Math Anxiety, Math Self-Concept, and Math Self-Efficacy in Adult Learners Compared to Traditional Undergraduate Students. *Adult Education Quarterly*, 64(4), pp.306-322. https://doi.org/10.1177/0741713614541461

Joyce, J., Hassall, T., Luis Arquero Montaño, J. and Donoso Anes, J.A. (2006). Communication apprehension and Mathematics anxiety as barriers to communication and numeracy skills development in accounting and business education, *Education + Training*, 48(6), pp.454-464. https://doi.org/10.1108/00400910610692967

Lee, C. and Johnston-Wilder, S. (2017). The Construct of Mathematical Resilience. In U. Xolocotzin Eligio (Ed.), *Understanding Emotions in Mathematical Thinking and Learning*. London: Academic Press (pp.269-291). https://doi.org/10.1016/B978-0-12-802218-4.00010-8

Marshall, E.M., Staddon, R.V., Wilson, D.A. and Mann, V.E. (2017). Addressing Mathematics anxiety within the curriculum. *MSOR Connections*, 15(3), pp.28-35. https://doi.org/10.21100/msor.v15i3.555

Morris, J. (1981). Math Anxiety: Teaching to avoid it. *The Mathematics Teacher*, 74(6), pp.413-417. https://doi.org/10.5951/MT.74.6.0413

Núñez-Peña, M.I., Suárez-Pellicioni, M. and Bono, R. (2013). Effects of math anxiety on student success in higher education. *International Journal of Educational Research*, 58, pp.36-43. https://doi.org/10.1016/j.ijer.2012.12.004

Núñez-Peña, M.I., Bono, R. and Suárez-Pellicioni, M. (2015). Feedback on students' performance: A possible way of reducing the negative effect of math anxiety in higher education. *International Journal of Educational Research*, 70, pp.80-87. https://doi.org/10.1016/j.ijer.2015.02.005

Núñez-Peña, M.I. and Bono, R. (2021). Math anxiety and perfectionistic concerns in multiple-choice assessment. *Assessment & Evaluation in Higher Education*, 46(6), pp.865-878. https://doi.org/10.1080/02602938.2020.1836120

Spencer, D. (2021). Understanding the coaching experiences of non-traditional students in Higher Education in the UK. *International Journal of Evidence Based Coaching and Mentoring*, 15, pp.84-95. https://doi.org/10.24384/00v3-nm61

Whitmore, J. (2009). *Coaching for Performance: GROWing Human Potential and Purpose: The Principles and Practice of Coaching and Leadership*, 4th ed. London: Nicholas Brealey.

Yaeger, D.S. and Dweck, C.S. (2020). What Can Be Learned From Growth Mindset Controversies? *American Psychologist*, 75(9), pp.1269-1284. https://doi.org/10.1037/amp0000794

Yan, V.X. and Schuetze, B.A. (2023). What is meant by "Growth Mindset"? Current theory, measurement practices, and empirical results leave much open to interpretation: Commentary on Macnamara and Burgoyne (2023) and Burnette et al. (2023). *Psychological Bulletin*, 149(3-4), pp.206-219. https://doi.org/10.1037/bul0000370

CASE STUDY

Zero waste: leveraging blended learning materials in traditional teaching

Cesare Giulio Ardito, Department of Mathematics, School of Natural Sciences, Faculty of Science & Engineering, University of Manchester, Manchester, UK.
Email: cesaregiulio.ardito@manchester.ac.uk.

Abstract

In recent years, and especially during the Covid-19 pandemic, several university course units have moved from a traditional format, made of face-to-face lectures supplemented by notes and exercise sheets, to a blended learning format based mainly on asynchronous engagement with recorded lectures, videos, and notes, followed by supplementary synchronous sessions to consolidate and expand on the core material. While during Covid-19 the latter format presented clear advantages, as things slowly go back to normal course leaders are free to choose between the two formats, each of which has its own set of advantages and challenges.

In this case study we explore the transition from a blended learning course back to a traditional face-to-face format, focusing on the use of an innovative 'anthology' approach to repurpose digital materials. Originally a blended learning course featuring a mix of videos, notes, and quizzes, the course underwent a transformation to adapt to a more conventional format without sacrificing the accumulated digital resources. Through this transition, an anthology of learning materials was created, which serves as supplementary support for student learning, providing a wealth of learning aids, ranging from mandatory activities to entirely optional exercises. Challenges encountered during this transition, specifically avoiding an increase in workload, and the maturity required from students, will be examined. We aim to provide a framework to highlight the potential of effectively reutilising blended learning resources within traditional teaching frameworks, while being mindful of workload issues, to enrich the students' learning experience.

Keywords: blended learning, anthology, digital resources, traditional teaching.

1. Introduction

In the wake of the Covid-19 pandemic, the educational landscape witnessed a transformative shift, characterised predominantly by the rise of blended learning (Anthony et al., 2022). This modality, an amalgamation of traditional synchronous sessions and online asynchronous educational resources, has not been merely a transient response to the pandemic, but rather emerged as a revolutionary method, here to stay. The allure of blended learning stems from its flexibility, its ability to cater to a diverse range of student needs, and the innovative potential it offers for course delivery. Nonetheless, as with any educational modality, it brings along some drawbacks when juxtaposed against traditional teaching methods, such as a reduced ability to build a personal relationship with students. Some educators still voice a preference for traditional teaching methodologies, citing superior outcomes and efficacy (Guppy et al., 2022), but the data is mixed. Further, we wonder: is traditional teaching genuinely superior, or do those educators find greater success with it simply due to their prolonged familiarity and decades of experience in employing it?

We examine a case of 'reverse' transition from a blended format to a traditional one. Our case study revolves around a 4th-year course, 'Noncommutative Algebra', offered at the University of Manchester (University of Manchester, 2022). The unit is comprised of a cohort that includes MMath and PhD students. Traditionally delivered face-to-face until 2019, the course adapted to a blended learning format in 2020, offering online asynchronous videos supplemented with review sessions and tutorials on Zoom. 2021 maintained the same approach, but with the addition of in-person synchronous 'review' sessions. After reviewing student feedback, the lecturer decided to return to the traditional mode of delivery in 2022. This oscillation prompted a compelling question: what becomes of the digital resources accumulated over the blended learning years?

In the following, we describe the implementation of an 'anthology' approach, a strategy designed to repurpose the vast reservoir of digital content without compromising the advantages of traditional delivery. Subsequently, we gauge the impact of this approach on student learning, assimilating feedback and outcomes to assess its efficacy. This is not meant as a formal statistical analysis, nor should be taken as such given the high potential for confounders represented by the several changes brought upon by the Covid-19 pandemic. Nevertheless, we describe the approach and provide a critical examination of the strengths and weaknesses of the process, with suggestions for any reader who wishes to apply a similar paradigm.

2. Implementation

2.1 Course format

The course 'Noncommutative Algebra' is a 4th-year advanced unit, rich in examples and diverse in its content, drawing from various mathematical areas previously encountered by students. Structured to facilitate an anthological approach, it aims to provide a broader perspective on the general theory through a detailed exploration of specific examples and fundamental noncommutative algebras such as quaternions, matrix rings over division rings, group representations and Weyl algebras.

Weekly, the course unfolds through two one-hour face-to-face lectures, complemented by a tutorial session, adhering to a traditional delivery format. Assessment within the course involves two take-home tests and a final closed-book exam, leaning heavily towards problem-solving, reflecting its advanced mathematical nature. This strategy aligns with the course's ethos of diving deep into noncommutative algebra's complexities, enabling students to progressively build up knowledge and expertise during the semester and then demonstrate their understanding effectively.

The course normally has approximately 25 students enrolled, half of which are normally 4th year MMath students or MSc students, with the other half made of PhD students in mathematics.

Blackboard, the chosen virtual learning environment, hosts the weekly learning plans and a multitude of resources, ensuring that materials are accessible, organised, and conducive to a consistent learning process.

2.2 The Anthology approach

We were able to make use of the 'leftover' resources from the blended learning implementations by adopting an 'anthology approach', made of an orchestrated interplay of various educational elements, each appropriately contextualised to enhance the students' learning experience within a traditional teaching framework, ensuring their effective and coherent integration to support and enrich the present course. To optimise the benefit for the reader, we choose to describe a refined approach, evolved from its 2022/23 iteration with minor modifications. This enhancement is informed by insightful student feedback, leading to modifications that enrich the overall strategy and execution. However, the core philosophy and elements of implementation remain the same.

The core element is the provision to each student of a meticulously crafted weekly plan, a structured guide curated with explicit learning objectives, providing students with a clear roadmap of the week's educational journey (see figure 1). It encompasses a full list of available content, paired with practical advice on managing workload, allowing students to navigate their weekly learning paths strategically. The plan is further enhanced with colour-coding of content, explicitly disclosed to students at the start of the course, serving as a visual guide that further facilitates students in organising their study time and focus effectively. The colours are structured as follows: core material (green), supplemental resources (orange), support channels (purple), and deeper dive (blue).

Week 1 (26th September) - Rings and algebras

In-person sessions:

- **Wednesday lecture:** Revision of background material. Definitions of rings, ideals, and algebras. Examples (matrices, group algebras, quaternions).
- **Friday lecture:** Important constructions on rings are revised/introduced, including subrings, ring homomorphisms, and product rings. Introduction of the Weyl algebra, one of the main examples in the course.
- **Tutorial:** Exercise sheet 2, exercises 2.1-2.6, 2.9-2.12 .

Course notes: From Section 2.1 up to and including Section 2.7, plus the definition of the Weyl algebra from 2.8.

Additional exercises:

- 2019-20 exam paper, exercise 2.
- Exercises (1-32) for chapter 0, Noncommutative Algebra, B. Farb, K. Dennis, Springer-Verlag, 1993.

Videos:

- Rings, subrings and ideals (32 minutes)
- The quaternions (7 minutes)
- Ring homomorphisms and product rings (13 minutes)
- Algebras (20 minutes)
- Endomorphism algebras and the definition of the Weyl algebra (17 minutes)

Additional support:

- Ask questions on our discussion board on Blackboard.
- Email me: cesaregiulio.ardito@manchester.ac.uk or come to my office hour.

Deeper dive:

- Group algebras (video - 23 minutes)
- Generalised quaternions: a visual approach (interactive web page).

Figure 2: Example of a weekly plan as displayed on Blackboard.

In line with a traditional delivery format, the plan emphasizes and identifies as ‘core activities’ the in-person review sessions and the tutorial. However, attendance is encouraged but not mandatory, allowing students the flexibility to choose engagement levels based on their individual learning needs and preferences. This flexibility ensures that the sessions are optimally beneficial, and attended by students who find real value and learning enhancement through these interactions.

The course is accompanied by a set of course notes, that form the ‘single point of truth’ for the main definitions and concepts and are meant for all students to be engaged with. Engagement with the ‘core’ activities, colour-coded as green, is sufficient to satisfy all intended learning objectives and to learn all the assessed content of the course.

The material is supplemented by concise, piecemeal videos, each ranging from 5 to 15 minutes and concentrating on distinct topics, colour-coded as orange. These videos mirror the lectures and the notes, acting as supplementary resources that provide additional perspectives and explanations and are meant to be engaged when additional support is beneficial for navigating more challenging subjects. Featuring a variety of academics beyond the course leaders, the videos introduce diverse teaching styles and explanations, offering students multiple approaches to understanding the material. The videos are independent of each other, allowing for flexibility and ease of reference, aligning precisely with the week's learning objectives or specific areas where students seek deeper understanding or clarification.

Exercises within the course are strategically organised into four distinct categories to foster a comprehensive and layered learning experience.

- Quick questions are seamlessly embedded within the course notes, serving as immediate touchpoints for reflection and consolidation of learning as students navigate through the course material (green).
- Traditional exercise sheets emulate the nature of exam questions, with a limited and manageable number, typically not exceeding ten each week. These exercises are instrumental for students to gauge their understanding and readiness for assessments (green).
- Complementary exercises provide an additional layer of engagement. Sourced from a variety of materials, they are designed for students seeking extra practice and reinforcement of concepts, either during the semester or the revision period (orange).
- Deeper dive exercises present an opportunity to explore beyond the conventional boundaries of the course. These challenges delve into advanced concepts or introduce novel notions not covered in the regular course notes. Specifically non-examinable, they cater to students with a heightened interest in a broader or more in-depth exploration of topics (blue).

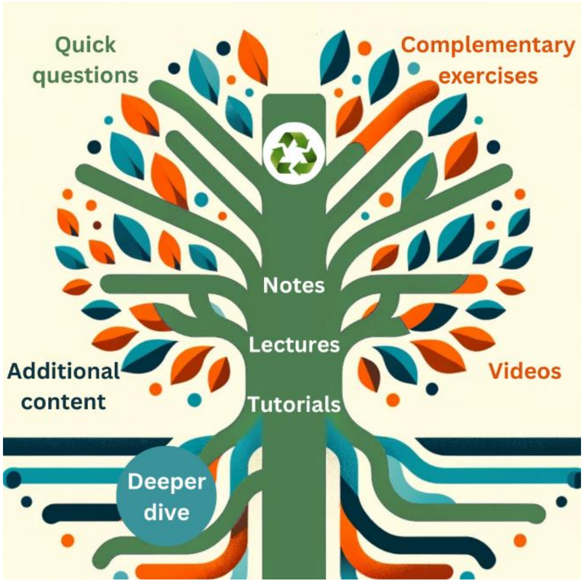


Figure 3: Visual diagram of the Anthology approach.

Furthermore, the ‘deeper dive’ section is enriched with additional content in a spectrum of formats, curated to guide and support students keen on delving into more advanced, additional topics outside the standard course outline.

Lastly, an online discussion forum is incorporated, fostering a space for collaborative learning, discussion, and peer interaction. This forum allows students to engage in discussions, seek and offer clarification, and explore various perspectives on the course content, enriching their learning experience. The forum is pseudonymised, meaning that each user is assigned an alias. Students are encouraged to participate and answer others’ questions, and staff moderate and occasionally intervene to answer more complex queries.

In summation, the Noncommutative Algebra course is structured to facilitate a comprehensive and adaptable week-by-week learning experience. The course interweaves various educational elements, such as in-person sessions, digital videos, and a range of exercises, aiming to offer a balanced and supportive learning environment. Each component is curated to allow students the flexibility to tailor their learning paths according to individual needs and preferences while preserving a common core. The anthology approach (see figure 2), marked by thoughtful organisation and an abundance of resources, serves to guide students through the course, providing a roadmap that aligns with their academic objectives and learning styles. This design aims to foster an environment conducive to focused and individualised exploration of noncommutative algebra, encouraging students to engage deeply with the material in a manner that resonates with their unique educational journey.

2.3 Impact

Before going into detail, it's essential to note that while the observed outcomes are positive, they cannot be exclusively or conclusively attributed to this approach alone. The unprecedented challenge of the Covid-19 pandemic brought sudden changes in student engagement and behaviour, which makes it challenging to definitively link any results to one specific factor or strategy.

A notable impact has been the augmentation of student engagement and participation with respect to previous years in the 2022-23 iteration. All 23 students consistently engaged with in-person activities. Further, a significant majority of students engaged with some of the supplemental materials provided within the course (see figure 3), suggesting that resources such as piecemeal videos and various exercise categories were found to be beneficial by the learners. Interestingly, there was no mention of workload issues in the end-of-year feedback survey, indicating that the course's design, with its emphasis on flexibility and student autonomy, effectively managed to avoid overwhelming the students.

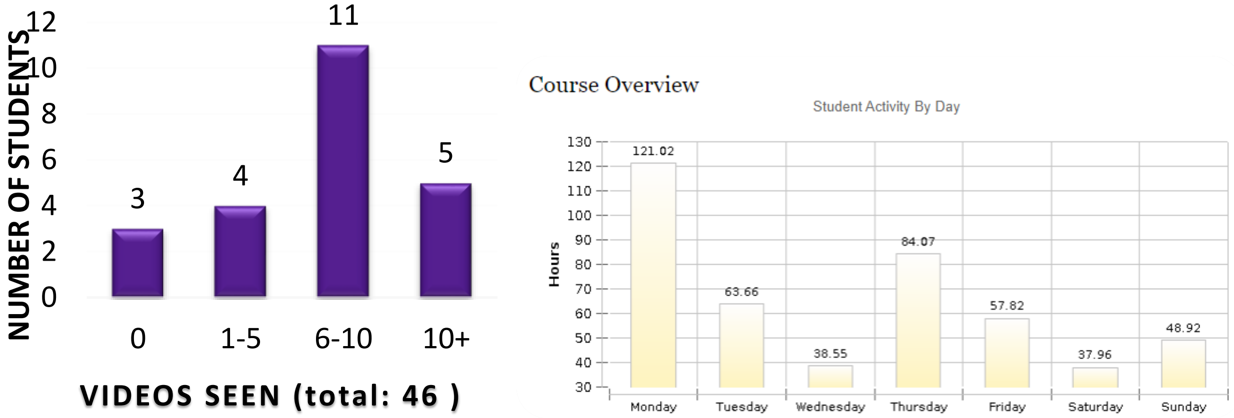


Figure 3: Students engagement with videos and with the Virtual Learning Environment

The course's design, which facilitated regular feedback loops, allowed for ongoing adjustments and refinements, ensuring that the course remained responsive to students' needs and experiences. This adaptability is a vital aspect of the course's impact, allowing for a continually evolving educational experience that aligns with each student cohort's needs and preferences.

Looking ahead, the experience suggests a promising trajectory for the course's future iterations. The insights gained from students' experiences and reflections provide valuable inputs that will inform and shape future adaptations and improvements in the course, ensuring that it remains aligned with students' needs and expectations.

Success rates in assessments have also seen improvement, indicating a potentially enhanced grasp of course material among students. This could be reflective of the course's ability to cater more effectively to different learning needs and styles, facilitating better preparation and understanding. A consequence of the multitude of resources available is also improved accessibility and flexibility, particularly beneficial for diverse learners, due to content being presented in different formats.

In terms of quantitative data, feedback from students and internal teaching evaluations have been predominantly positive. Over 90% of students reported high satisfaction levels, which speaks volumes about the course's effectiveness and its ability to meet student needs and expectations. The course received a User Experience Questionnaire (UEQ) score of 4.8/5, reflecting a high level of student satisfaction and positive user experience. This score, coupled with the qualitative feedback, paints a picture of a course that has been well-received by students, managing to meet their educational needs effectively while providing a positive and supportive learning environment.

3. Discussion

The implementation of the anthology approach in the Noncommutative Algebra course unfolded on what can be considered an ideal testing ground. Comprising 4th-year and PhD students, the course was populated by experienced learners, well-versed in independent learning and efficient time management. These attributes facilitated a conducive environment for the anthology approach to be employed and evaluated. Moreover, the course's intrinsic structure, characterised as an anthological collection from a broader mathematical tapestry, naturally aligned with the approach, allowing for seamless integration of 'deeper dive' content and other enriched materials.

A significant asset in this implementation was the existing 'piecemeal' format of the videos, as mentioned above. This modular construction allowed for effortless adaptation and customisation of the content, aligning with the evolving needs and focuses of the course (Thompson et al., 2021). Modern resources, such as Canva (https://www.canva.com/), proved invaluable in this aspect, enabling easy editing and modification of video content and minimising the amount of technical work needed for the lecturer to implement this approach.

The anthology approach's implementation did not necessitate an exhaustive investment of time or resources. Its design allowed for the flexible inclusion of various materials, even as optional content, ensuring that the course could remain responsive to changing needs or focuses without requiring extensive modifications. We highlight that the approach is suitable even when only a limited amount of 'deeper dive' material is available for distribution.

Suitability emerges as a key consideration in the discussion. The course's advanced level, combined with the experienced learner demographic, was identified as particularly congruent with the anthology approach. This suitability suggests a potential direction for the consideration of this approach in future course designs, highlighting advanced courses with experienced learners as particularly promising arenas for its application and evaluation.

Incorporating insights from the course's execution, the approach reveals a pattern of positive impacts, such as increased student engagement and successful assessments. However, it also brings forth areas for consideration and refinement, such as the diversification of resource formats and the continual gathering of evidence to assess and enhance the approach's effectiveness. These reflections form a basis for ongoing improvement, ensuring that the approach remains dynamic and responsive in meeting the educational objectives and needs of future courses.

4. Aspects to consider for the future

Implementing the anthology approach in courses with less experienced students, such as those in their first year, would necessitate thoughtful considerations about potential workload challenges (Meehan and McCallig, 2019). The substantial volume of materials, while rich in content and diversity, could potentially overwhelm students who may not yet have honed the skills necessary for efficient and critical engagement with extensive resources. There's a risk of these students feeling overwhelmed if they attempt to engage uncritically with the entirety of the material available, without a strategic approach to prioritise their engagement based on relevance and individual learning needs (Phillips, Schumacher and Arif, 2016; Banihashem et al., 2023). On the other hand, the abundance of available pre-recorded material may also reduce attendance for in-person activities, which can be negatively correlated with academic performance (Trenholm et al., 2019).

Another aspect to consider is the temporal fixation of video content. This permanence can pose challenges when updates or changes in the course are needed, as the videos may not easily allow for modifications or adaptations in line with the evolving curriculum or new pedagogical insights.

Keeping track of a vast array of materials also emerges as a potential area requiring attention. This necessitates an ongoing commitment to oversight and coordination to ensure that the contents remain relevant, updated, and conducive to the course's learning objectives and overall coherence.

Internal coherence within the course is paramount. Careful attention must be given to maintain consistency, such as in notation and the presentation of concepts, to ensure that the course unfolds as a cohesive and integrated learning journey. This is essential to avoid confusion and to support students in building a structured and coherent understanding of the subject matter.

Lastly, while the approach brings flexibility and richness of content, it may also introduce constraints when considering substantial course revisions or shifts in direction. The presence of an extensive array of existing materials may inadvertently influence or limit the scope and direction of future course enhancements and redesigns.

In conclusion, while the anthology approach presents numerous advantages and has demonstrated preliminary positive impacts in its application, a thoughtful and critical evaluation of these aspects is essential to navigate its future implementations successfully, ensuring that it continues to serve as an effective and supportive educational strategy, and it may not ultimately be suited for all courses in higher education.

5. References

Anthony, B. et al. (2022). Blended Learning Adoption and Implementation in Higher Education: A Theoretical and Systematic Review. *Technology, Knowledge and Learning*, 27(2), pp.531-578. https://doi.org/10.1007/s10758-020-09477-z

Banihashem, S.K. et al. (2023). Modeling teachers' and students' attitudes, emotions, and perceptions in blended education: Towards post-pandemic education. *International Journal of Management Education*, 21(2), Article 100803. https://doi.org/10.1016/j.ijme.2023.100803

Guppy, N. et al. (2022). The post-COVID-19 future of digital learning in higher education: Views from educators, students, and other professionals in six countries. *British Journal of Educational Technology*, 53(6), pp.1750-1765. https://doi.org/10.1111/bjet.13212

Meehan, M. and McCallig, J. (2019). Effects on learning of time spent by university students attending lectures and/or watching online videos. *Journal of Computer Assisted Learning*, 35(2), pp.283-293. https://doi.org/10.1111/jcal.12329

Phillips, J.A., Schumacher, C. and Arif, S. (2016). Time spent, workload, and student and faculty perceptions in a blended learning environment. *American Journal of Pharmaceutical Education*, 80(6), Article 102. https://doi.org/10.5688/ajpe806102

Thompson, P. et al. (2021). The Effect of Designing and Segmenting Instructional Video. *Journal of Information Technology Education: Research*, 20, pp.173-200. https://doi.org/10.28945/4756

Trenholm, S. et al. (2019). Investigating undergraduate mathematics learners' cognitive engagement with recorded lecture videos. *International Journal of Mathematical Education in Science and Technology*, 50(1), pp.3-24. https://doi.org/10.1080/0020739X.2018.1458339

University of Manchester (2022). *Noncommutative Algebra - Course unit details*. Available at: https://www.manchester.ac.uk/study/undergraduate/courses/2023/01688/mmath-mathematics/course-details/MATH42042#course-unit-details [Accessed: 3 November 2023].