RESEARCH ARTICLE

Making Mathematicians: Developing expert attitudes with authentic maths activities.

Matthew Jones, School of Management, UCL, London, U.K. Email: matthew.m.jones@ucl.ac.uk
Zainab Kazim Ali, Design Engineering and Mathematics, Middlesex University, London, U.K. Email: z.kazim@mdx.ac.uk
Snezana Lawrence, Design Engineering and Mathematics, Middlesex University, London, U.K. Email: s.lawrence@mdx.ac.uk
Brendan Masterson, Design Engineering and Mathematics, Middlesex University, London, U.K. Email: b.masterson@mdx.ac.uk
Alison Megeney, Design Engineering and Mathematics, Middlesex University, London, U.K. Email: a.megeney@mdx.ac.uk
Nicholas Sharples, Design Engineering and Mathematics, Middlesex University, London, U.K. Email: n.sharples@mdx.ac.uk

Abstract

One of the goals of an undergraduate degree in mathematics is to transform students' perceptions of mathematics from calculations with the rote application of formula to the reflective, creative problem-solving that is highly valued in academia and other professions. This can be achieved by incorporating authentic mathematical activities (i.e. the kind of tasks a maths graduate can expect in the workplace) into the design and delivery of undergraduate programmes. The Middlesex maths team have implemented a variety of novel teaching and learning methods into their specialist maths provision to achieve this aim. Our approach includes the use of generative artificial intelligence; extended, vague, problem-solving assignments; student choice in assessment; and reflective components. In this paper we discuss the implementation, benefits, and challenges of these authentic mathematical activities, focusing on their effect on students' perceptions of mathematics during their studies. We use questionnaires to determine how students' perceptions of mathematics change while doing these activities and their attitudes to the activities themselves.

Keywords: Authentic assessment, attitudes, perceptions, identity, Mathematics Attitudes and Perceptions Survey, problem-solving, choice in assessment, artificial intelligence, reflection.

1. Introduction

Authentic assessment in higher education can be characterised as "assessment requiring students to use the same competencies, or combinations of knowledge, skills and attitudes that they need to apply… in professional life" (Gulikers, Bastiaens and Kirschner, 2004) and is a significant shift from traditional exam-based assessment, which is itself highly dissimilar to professional practice. In a systematic review Villarroel et al. (2018) conclude that authentic assessment has a positive impact on a variety of abilities related to employability, such as autonomy, motivation and self-regulation. Further, the authors distil authenticity into three dimensions: realism, cognitive challenge and evaluative judgement, and provide a framework for designing and operating authenticity assessment in higher education programmes.

In professional life, mathematics graduates have a reputation for analytical thinking and problem-solving that are particularly "sought after by employers" (QAA, 2023). One goal of authentic assessment in mathematics, therefore, is to ensure students explicitly develop these skills as part of their mathematics undergraduate experience, which requires more from them than simply being able to perform well in controlled exam conditions.

Authentic assessment has been a characteristic of Middlesex University mathematics degrees, with reflection, communication and problem-solving activities included in the degree programme when it relaunched in 2014 (Megeney 2016). The learning and teaching strategy for these programmes has been continually developed: all exams (high-stakes, controlled environment, end-of-module summative assessment) were removed in September 2021 (although some mid-module, low-stakes in-class tests remain) in favour of authentic assessment, which attempts to emulate the specialist work that mathematics graduates will perform in their professional roles. In Masterson et al. (2024) we distinguish between "authentic assessment", whose outputs have professional analogues (such as reports, computer code, or presentations) and "authentic problems" whose inputs are typical of professional environments (for example vague, imprecise, or requiring a significant element of judgement to begin).

In this paper, we examine the effects of authentic assessment on student attitudes and perceptions of mathematics. The premise is that if students are persistent and confident in mathematics, believe in its applicability to the real world, and are learning to develop understanding rather than just for completing tasks then they have expert-aligned attitudes that they can apply in professional life. We examine this relationship by deploying the Mathematical Attitudes and Perceptions Survey (MAPS) on undergraduate mathematics students whose degree programmes contain many authentic activities.

In the next section we describe four categories of authentic activity (*Problem-solving, Artificial intelligence, Reflection, and Choice in assessment*) that are features of the Middlesex university undergraduate degree programmes (two further categories, *Communication* and *Outreach* are described in Jones et al. 2025). First, we introduce the MAPS survey tool.

1.1. The Mathematics Attitudes and Perceptions Survey

The Mathematics Attitudes and Perceptions Survey (MAPS) was developed by Code et al. (2016) to characterise undergraduate student perceptions of mathematics in educational settings. The survey consists of 32 statements which students respond to with a 5-point Likert scale (from "Strongly Disagree" to "Strongly Agree"). It contains seven subscales: *Confidence* (the perceived ability to successfully engage in mathematical tasks), *Growth mindset* (the belief that mathematical ability can be developed rather than an innate property), *Real world* (the view on the applicability of mathematics to everyday life), *Persistence* (the attitudes and approach when getting "stuck" on a problem), "Interest" (the motivation in studying mathematics), *Sense making* (learning mathematics for understanding, rather than for completing tasks), and *Answers* (on what form solutions to mathematical problems can take). The authors performed a factor analysis on a pool of specialist and non-specialist maths students in North American universities ($N = 3411$). The whole instrument achieved a Cronbach's alpha of 0.87 (95% confidence interval $[0.86, 0.88]$). An expert consensus, determined from a panel of mathematics faculty at British Columbia University ($N = 36$), was obtained for 29 of the MAPS statements. The MAPS instrument therefore quantifies the extent to which undergraduates' attitudes and perceptions of mathematics agrees with that of experts.

In their initial study Code et al. (2016) found that students' attitudes varied across courses, with first-year Calculus 1 courses having students with lower MAPS scores (i.e. attitudes towards mathematics less aligned with those of experts) compared with students on a second-year Introduction to Proof course. They also reported a generally positive correlation between student grades and their MAPS scores. A further study by Maciejewski et al. (2021) demonstrated lower

MAPS scores for first-year students taking “development” maths courses (required, but non-credit bearing, courses for students entering university with low mathematics attainment) compared with first-year students taking “college-level” maths courses for credit.

Of interest to the present work, Code at al. (2016) also examined students who completed the survey in both the September and April of their first academic year, reporting “All MAPS categories… saw declines over the academic year”, which is to say that students’ attitudes to mathematics moved further away from those of expert mathematicians during their first year at university. Further, this decline was present even for courses using “flipped” instructional methods (although the decline in the *Real world*, *Persistence*, *Sense making* and *Answers* categories weren’t statistically significant). The authors speculate that this is partially due to first-year courses’ emphasis on “solving low-level inauthentic problems”. A similar first-year decline in expert-aligned attitudes towards physics (using a related survey) is reported in Cahill et al. (2014).

In contrast, Ozimek et al. (2024) report that amongst prelicensure nursing students, those undertaking a first-year Clinical Maths module designed with “problem solving in authentic situations” have greater MAPS scores compared with those who have yet to take the course. These results suggest that authenticity is an important element of maths courses that aim to develop students into expert mathematicians.

2. Authentic activities

2.1. Problem-solving

The QAA (2023) recognise “the fundamental nature of MSOR as a problem-based subject area” and emphasise that graduates should have skills “in the solution of new problems arising in professional work or in further study”. Problem solving ability has traditionally been assessed in the second half of exam papers through applications of familiar theory to previously unseen, but perhaps not entirely unfamiliar problems. In our view this kind of problem solving is inauthentic for two reasons: first, mathematics professionals are more likely to identify or develop appropriate theory in response to specific problems, rather than look for applications of a particular theory. Second, problem solving in the professional domain involves a wide range of skills such as research, experimentation, collaboration and communication that cannot be assessed in controlled exam conditions.

Since its inception, the Middlesex mathematics degree has included a core level 5 module (Year 2 of an undergraduate degree in the UK), dedicated to developing the desirable problem-solving skills mentioned above (Jones and Megeney 2019, and Masterson et al. 2023). Following revalidation in 2022, the module increasingly emphasises the “mathematization” (see Freudenthal 1968) of vague and imprecise problems rather than solving a problem already expressed mathematically. For example, the (space-themed) 2023-24 assessment consisted of 8 possible problems, including

1. A solar powered satellite is on an elliptical orbit of the sun. Explore how much energy the satellite can generated per orbit and per hour. What capacity batteries should the satellite carry?
2. Suppose an object is moving around randomly in space. Will it ever reach Earth?
3. Celestial bodies tend to be approximately spherical. Can we measure how spherical they are in order to compare them?

Small groups of students chose a problem to work on over an eight-week period with the brief that they are “primarily being assessed on the approach [they] take to solving problems, rather than the actual solution.” Indeed, only 10% of marks were allocated for the “solution” of the problem, with the more marks available for an account of the “problem solving” process (10% for an introduction, mathematical formulation and strategy, and 30% for a thorough description of the problem-solving process, including which activities did and did not help the group make progress, together with an account of how the group interacted). At the start of the module, students learn problem-solving approaches such as Polya (1990), Mason et al. (2010) and Bransford and Stein (1993) to develop their problem-solving skills and give them the language and choice of frameworks to express and reflect on their process (Jones and Megeney, 2019 and Masterson et al., 2023).

Vague, open-ended problem solving has also been incorporated into other modules such as level 5 *Mathematical Statistics*, where students have an unconstrained choice of dataset to analyse and must identify an interesting question, choose an appropriate analysis to perform, and are given agency in demonstrating they have satisfied the learning objectives (Masterson et al. 2024).

2.2. Artificial intelligence

In November 2022 ubiquitous access to generative artificial intelligence began with the release of ChatGPT-3.5 by OpenAI. Universities recognised risks in this technology, particularly concerning ethical use, equality of access, and academic integrity but also opportunities in transforming learning experiences and increasing graduates’ employability and productivity through AI literacy (Russell Group, 2023). Responses have been mixed, with some courses returning to the pre-covid practice of controlled assessment conditions (such as exams), while others explicitly include the use of generative AI as an option or requirement in assessment.

Whether AI can generate novel mathematical research remains to be seen. However, eminent researchers have recognised that potentially “it may only take one or two further iterations [for] the tool being of significant use in research level tasks” (Tao, 2024). It is important, therefore, for students to experience using generative AI tools in order to understand the capabilities, limitations, and appropriate uses of this technology for mathematical activities. In any case, following the reported productivity gains of using large language models for workplace tasks (Dell’Acqua et al. 2023), universities should prepare students for careers in which these technologies are widely used for knowledge work and decision-making. The think-tank Demos suggests that the challenge for universities in this setting is to “equip their graduates with skills, competencies and dispositions that will enable them to offer something different in the workplace” (Demos, 2023).

In response, the Middlesex mathematics team have begun including generative AI into the syllabus and assessment of their mathematics programmes. The aim is that as students develop their mathematical and statistical knowledge they also develop the AI literacy to be judicious in its use, critical of its output and are able to articulate the value of their knowledge and skills in comparison to a naïve user of generative AI.

In assessment for the level 5 *Mathematical Statistics* module, students first complete some data analysis questions then write a reflective report comparing their solutions to attempts made by a large language model (Masterson et al. 2024). In the final piece of assessment students will be invited to incorporate generative AI into their data analysis directly, but explicitly.

Students are also permitted to use generative AI to support their extended written communication tasks in the level 5 module *Problem-Solving and Communication*. These briefs are centred on experiences that the students have had such as museum trips or guest lectures (Jones et al. 2025), and students are strongly encouraged to personalise their work and write from their own perspectives. Consequently, although students might use generative AI to help structure or redraft work, the content and the judgement of what to include, will be their own.

In the *Problem-Solving and Communication* syllabus, generative AI is also used to support students in producing high-quality mathematical animations using the manim library (see, for example, Sanderson 2024). Writing manim code requires a non-trivial understanding of object-oriented Python, which ordinarily has a significant learning curve. However, after a single two-hour workshop using generative AI to write manim code, students were able to produce high-quality animations of their own design, which they could optionally use as part of an assessment.

2.3. Reflection

Villarroel et al. (2018) identify evaluative judgement as an important dimension of authentic assessment. The aim is for students to "develop criteria and standards about what a good performance means" ultimately endowing the student with "the lifelong capacity to assess and regulate their learning and performance", which will enhance their employability.

At Middlesex university, evaluative judgement is developed in reflection activities that are embedded throughout the mathematics programmes. In the level 4 (Year 1 of an undergraduate degree in the UK) module *Data and Information*, students reflect on their participation in outreach activities (such as SMASHfest see Griffiths and Keith 2021, Jones et al. 2025 and Megeney 2016). These reflections focus primarily on their experience of interacting with staff, fellow students and the public, and how these experiences may benefit their mathematics education.

At level 5 (Year 2 of an undergraduate degree in the UK), the module *Problem-Solving and Communication* is assessed through group problem-solving (discussed in Section 2.1) and a portfolio of communication briefs (Jones et al. 2025). Students also submit a written reflection on their assessment (worth 15% of their overall grade) that requires them to comment on their group dynamics, evidence their improvement as problem-solvers, and critically evaluate both their submitted communication briefs and the process they followed to produce them. At this level, students are given a structured set of prompts to support a high-quality reflection. Students are asked to consider the knowledge and skills (including time management) that they have used; to identify particular challenges; to justify their approach; to explain how they engaged with formative feedback; and describe what could be improved or would be done differently in the future. These reflective questions have natural synergies with problem-solving assessment as Polya's (1990) fourth step is "looking back", where students focus on checking the result and argument, and seeing if the result can be derived differently, or is now immediately clear from the student's new point of view.

At level 6 (Final year of an undergraduate degree in the UK), reflective tasks are present but are less explicitly structured so that students can take ownership of this process. The *Project* module, for example, requires students to keep a reflective diary of meetings with their supervisor, but only gives minimal guidance on the contents. The module *Real and Complex Analysis* builds on the students' reflective ability by including evaluative judgements of their peer's work.

2.4. Choice in assessment

Student engagement in learning can be improved by granting a degree of agency over their assessment as "flexibility in assessment allows students to take a proactive role in their learning" (Pretorius, van Mourik and Barratt, 2017). The Middlesex maths team have incorporated student choice into assessment in two distinct ways: first, we allow students to select the format of their submission, which could be written, drawn, audio recorded, or video recorded, or a combination of these formats for an individual piece of coursework. This choice of format was present in the level 5 *Problem-Solving and Communication* module, as well as more mathematically technical modules such as the level 6 *Real and Complex Analysis* where, for example, students choose the format of their proof that multivariable polynomials are differentiable.

Student choice of assessment format is facilitated in part by the three-year loan of iPads to all undergraduate maths students (Jones, Megeney and Sharples 2022), which provides a common, equitable platform for multi-media submissions. Students may wish to choose the format that they believe best demonstrates their mathematical ability or take the opportunity to develop their skill and confidence in other formats. Including this wide choice made the assessment more accessible, reducing the need for reasonable adjustments and improving the inclusivity of the modules.

Second, in many module assessments students have a substantive choice over the question they answer (Masterson et al. 2024). This includes a free choice of dataset and method of analysis in the level 5 *Mathematical Statistics* module (with some guidance on the learning outcomes expected to be demonstrated, such as the calculation and interpretation of confidence intervals), and a choice of vague problem for group work in level 5 *Problem-Solving and Communication* (see Section 2.1).

These significant choices in assessment require students to have a good overview of the subject and to be able to make sound judgements on their overall approach to a task, which are desirable features for graduate mathematicians. However, there is a risk that students make unsuitable choices, for example on a topic that is particularly challenging to communicate in an audio-only format, such as geometry, or a dataset that isn't suited to the learning outcomes, for example categorical data for the calculation of confidence intervals. However, formative submissions, and frequent check-ins serve as opportunities to guide students' judgement and reinforce their learning to correct any such errors before summative submission.

3. Methodology

We conducted an anonymous survey between July and August of 2024 of current students and recent graduates of our undergraduate mathematics programmes BSc Mathematics, BSc Mathematics with Computing, and BSc Mathematics and Data Science. The survey was adapted from the MAPS in the following way: for brevity the three statements in the *Interest* subcategory and the two uncategorised statements were removed. Our rationale for this decision was that the MAPS was designed for undergraduates taking general maths modules, whereas students who had chosen to study specialist mathematics degrees had an established interest in mathematics. We also introduced the statement "I am a mathematician" to explore students' feelings of mathematical identity, giving a total of 28 statements.

In the survey, participants were first asked to recall how they felt about mathematics before joining Middlesex University and record their recollection through rating their agreement with the MAPS and identity statements. Next, for each of the six learning and teaching elements *Problem-solving, Artificial intelligence, Reflection, Choice in assessment, Communication, and Outreach* participants were asked whether they'd encountered this element: an affirmative response gave four additional statements "[element] made me more confident in mathematics", "[element] increased my anxiety", "[element] helped me prepare for my professional career" and "[element] is an important part of mathematics", and a free-text response about their experience of this element. Finally, participants were given the MAPS and identity statements a second time to rate how they currently felt. The survey and research methodology were approved by the Middlesex University ethics committee ref: 28902.

Participants rated each statement on the scale *Strongly disagree, Disagree, Neither agree nor disagree, Agree,* and *Strongly agree*. Our analysis differs from that of Code et al. (2016) who dichotomise responses into scoring one point if the direction of the participants response (agree or disagree) matches that of the expert consensus, and zero points otherwise. This gives an easily interpretable "consensuality" measure similar to the common positivity measure used for Likert data (Jeong and Lee, 2016), which is statistically robust as it models agreement with each statement as a Bernoulli random variable.

Instead, to also capture the strength of agreement, we score from -2 (Strongly disagree) to 2 (Strongly agree) before multiplying by -1 if necessary to align the scale with the direction of the expert consensus for that statement. We then calculate an average score for each participant, question category and time. This "intervalist" analysis of Likert data is regarded as statistically robust (Carifio and Perla, 2008, Norman 2010, Sullivan and Artino, 2013) although there has been substantial debate about the appropriate analysis of attitude survey data since Likert (1932) introduced his eponymous methodology. An advantage of the approach taken here is that small changes in response (for example from *Agree* to *Strongly agree*, or from *Neither agree nor disagree* to *Disagree*) are captured in the average scores. We report summary statistics across these averages in the next section.

The non-MAPS statement "I am a mathematician" and the statements about the six specific teaching and learning elements do not have an established expert consensus, so are simply scored from -2 (*Strongly disagree*) to 2 (*Strong agree*), with the exception of "[element] increased my anxiety" for which this scale is reversed. Consequently, we report data so that positive scores for MAPS statements indicate alignment with the expert consensus, and positive scores for other statements have a positive connotation (more confidence, less anxiety, etc.).

As the survey was administered at a single time point our methodology isn't a true pre/post study to measure change in attitudes, which is a limitation of the design. Participants may be biased in their recollection of pre-university attitudes, perhaps by judging prior attitudes more harshly in light of their current mathematical experience. Our "quasi-pre/post" design, however, does capture the extent to which participants currently perceive how their attitudes have changed since before joining university.

Data analysis was performed in the R programming language and the code is available in Sharples (2025). In this paper we report on the MAPS and identity statements, and participant perceptions about *Problem-solving, Artificial intelligence, Reflection,* and *Choice in assessment*. For discussion and results on the *Communication* and *Outreach* activities see Jones et al. (2025).

4. Results

A total of $N = 13$ surveys were completed. Participants reported having joined the university between 2014 and 2021 so cover a wide range of the Middlesex mathematics provision. We also include the results from an additional participant who completed only the first half of the survey. All other surveys were fully completed apart from one missing response for question 6 at the "before university" time point.

4.1. MAPS and identity

Overall results are summarised in Table 1: we see that, on average, participants' attitudes before joining university to study a specialist mathematics undergraduate degree mildly align with the expert consensus (mean score of 0.554), but there is an improvement in their attitudes after completing at least some undergraduate study (mean score of 0.899). Further, before university the least expert-like participant disagreed with the expert consensus overall (mean score of -0.24), but after some undergraduate study all students agreed with the consensus overall (minimum mean score of 0.423).

There was a substantial shift in the attitudes about mathematical identity. Recollecting their attitudes before joining an undergraduate mathematics degree, of the 14 participants 6 agreed and 4 strongly agreed with the statement that they were mathematicians. At the time of the survey this improved to 5 agreeing and 8 strongly agreeing out of 13 respondents. The average Likert scores can be seen in Table 1.

Table 1: Likert scores comparing the recollection of attitudes before university (pre) and attitudes at the time of completing the survey (post). Minimum and maximum are taken across the average scores for each participant.

Questions	Time	n	Mean Likert score	sd	Min	Max
MAPS subset	pre	14	0.554	0.443	-0.24	1.08
MAPS subset	post	13	0.899	0.375	0.423	1.69
Identity	pre	14	0.857	1.03	-1	2
Identity	post	13	1.62	0.506	1	2

Results for individual question categories can be seen in Figure 1. For each category there was a closer alignment of attitudes to the expert consensus following some undergraduate study, particularly in the *Persistence* (+0.49), *Confidence* (+0.47), and *Answers* (+0.43) categories.

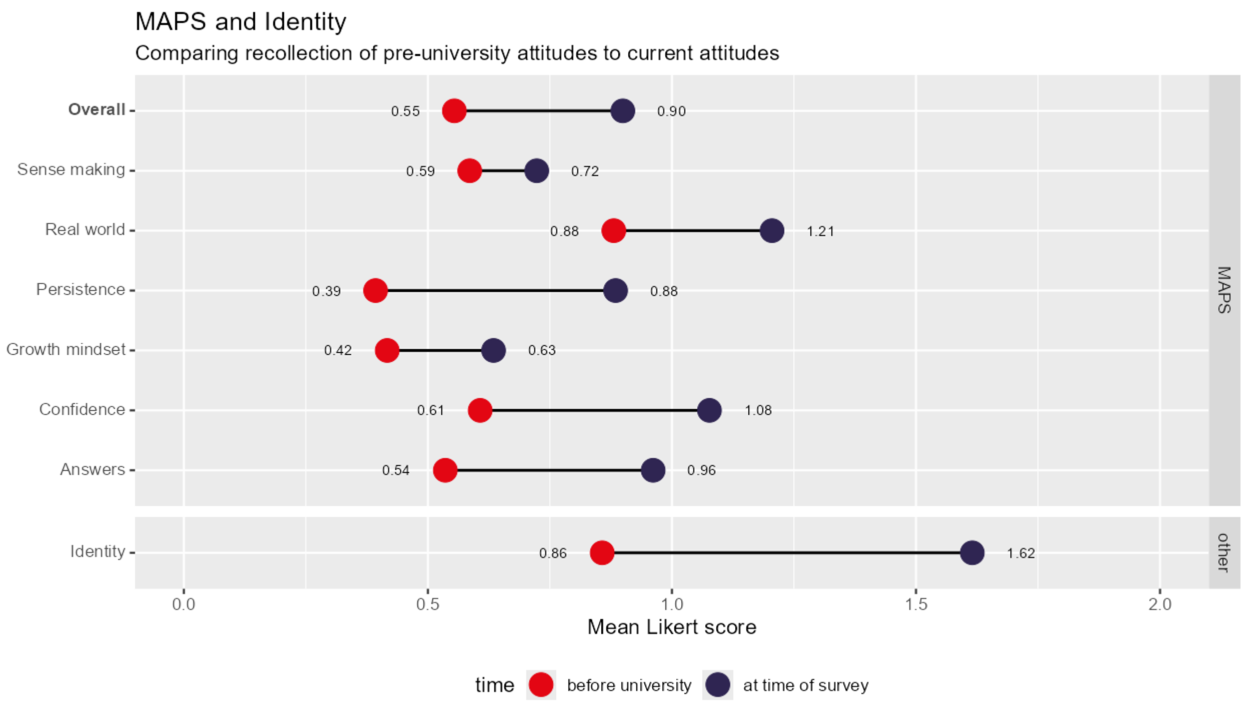

Figure 1: Comparison of students' attitudes aggregated by MAPS category (positive scores show agreement with the expert consensus) with an additional question on mathematical identity.

4.2. Attitudes to teaching and learning elements

In general, respondents had positive attitudes about the use of *Artificial intelligence, Choice in assessment, Problem-solving,* and *Reflection* in their degree programmes (see Table 2). There was substantial agreement that all these teaching elements improved respondent confidence, and that they are important to mathematics. Further, respondents substantially agreed that *Problem-solving* helped prepare them for careers, but felt less strongly that the other teaching elements supported this outcome. Respondents agreed that *Problem-solving* and *Choice in assessment* didn't increase anxiety, but were more ambivalent about anxiety from the *Artificial intelligence* and *Reflection* elements.

Table 2: Likert scores of respondent attitudes to selected teaching and learning elements.

Element	Statement	n	Mean Likert score	sd
Artificial intelligence	Improved confidence	4	1.25	0.5
Artificial intelligence	Important to mathematics	4	1.75	0.5
Artificial intelligence	Career preparation	4	0.75	1.26
Artificial intelligence	(didn't) increase anxiety.	4	0.25	1.5
Choice in assessment	Improved confidence	14	1.29	0.73
Choice in assessment	Important to mathematics	14	1.21	0.70
Choice in assessment	Career preparation	14	0.50	1.09
Choice in assessment	(didn't) increase anxiety.	14	1.29	0.61
Problem-solving	Improved confidence	14	1.71	0.47
Problem-solving	Important to mathematics	14	1.86	0.36
Problem-solving	Career preparation	14	1.5	0.85
Problem-solving	(didn't) increase anxiety.	14	1.21	0.89
Reflection	Improved confidence	11	1.45	0.52
Reflection	Important to mathematics	11	1.18	1.17
Reflection	Career preparation	11	0.91	1.04
Reflection	(didn't) increase anxiety.	11	0.64	1.36

Looking at the individual responses in Figure 2, we note that attitudes to *Problem*-solving were generally positive, while attitudes to *Choice in assessment* were mainly positive and neutral, whereas both *Artificial intelligence* and *Reflection* polarised respondents. In particular, some students stated that the use of *Artificial intelligence* and *Reflection* increased their anxiety and that *Reflection, Choice in assessment*, and *Artificial intelligence* didn't help prepare them for their careers.

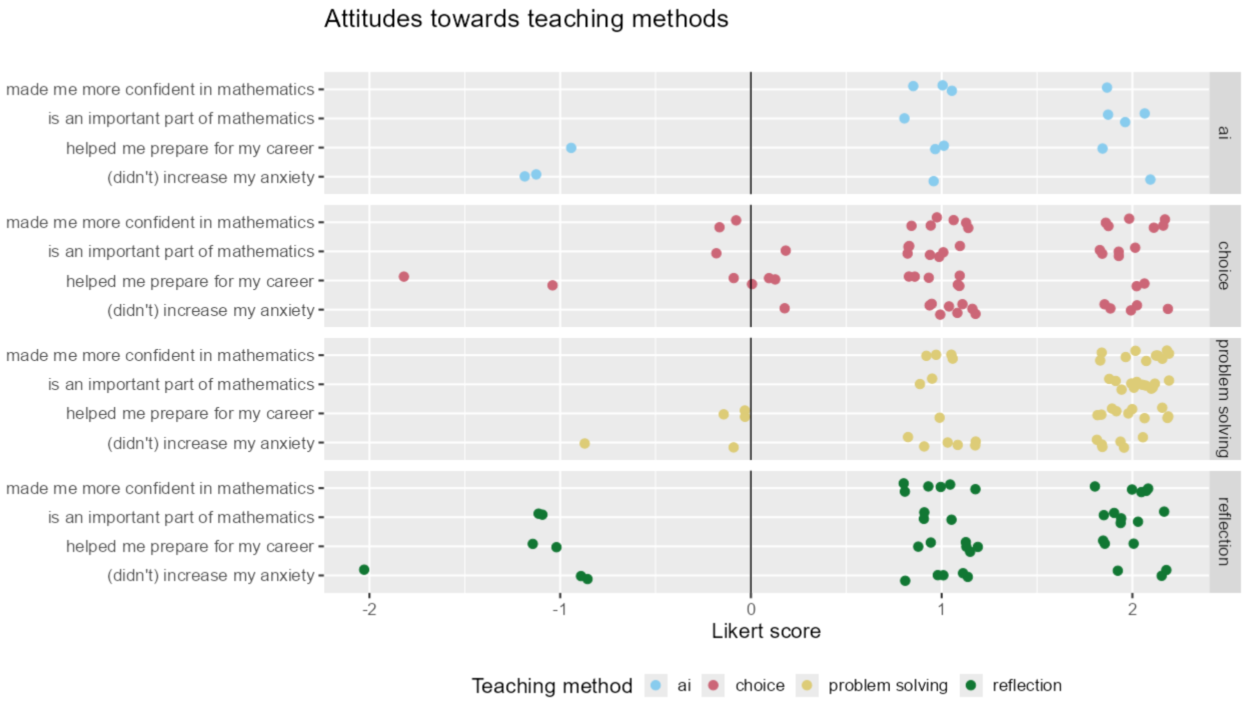

Figure 2: Comparison of attitudes to selected teaching and learning elements. Each point represents an individual respondent's Likert rating. Points are jittered to prevent overplotting.

Respondents who left free text comments were positive about the teaching elements. They reported benefits to individual approaches to learning as *Choice in assessment* "allowed me to properly understand my capabilities and plan my studying accordingly" and "is helpful as we all have different ways of learning and interests".

Benefits to learning outcomes were also reported as *Problem-solving* "requires time and development [but]… leads to students thinking independently" and "allowed me to think more outside of the box when it came to answering questions", while *Choice in assessment* "helps [in learning] more about a specific approach to a problem" and *Reflection* "is an important part of mathematics learning… [that leads to] improvement in the area [reflected upon]".

Further, the free text comments were used to express appreciation with respondents "grateful to have had*" Choice in assessment* and reporting that they "Loved the problem-solving module!". Respondents also gave suggestions for future developments reporting that *Choice in assessment* would "give students more flexibility and time management" and that *Artificial intelligence* "should be explored and used by future undergraduate students and lecturers… [It] can revolutionize the education of mathematics".

5. Discussion and future work

The Middlesex university mathematics degree programmes have been designed with authentic activities embedded throughout, to develop graduate skills and therefore employability in students (Megeney 2016). Students on these programmes report that their attitudes to mathematics have substantially shifted towards those of experts, compared to the recollection of their pre-university attitudes. Although we didn't take an objective measure of the student's attitudes at the start of their course, at least at the time of the survey the students believed that they had become more expert-like in their attitudes.

It is plausible that the inclusion of authentic activities is at least partially responsible for the expert-alignment of students' attitudes for two reasons: first, the students themselves generally report that these activities improved their confidence, helped them prepare for their careers, and are important parts of mathematics. Second, other studies using the MAPS show attitudes becoming *less* expert-like in both traditional and flipped classroom mathematics courses with assessment "depending largely on traditional written exams in all cases" (Code et al., 2016). This is perhaps suggestive that inauthentic exams have a detrimental effect on student's mathematics attitudes, which is in line with the known negative effects of exams in higher education in general (French, Dickerson and Mulder, 2023).

A larger scale study, with pre and post MAPS measures, across multiple universities with differing maths provisions would be necessary to robustly establish if the inclusion of authentic activities such as those described above has a significant effect on student attitudes. However, the impact of this work is potentially very high as these attitudes include those that employers recognise as being particularly sought after in graduates.

Generative AI activities are a recent inclusion in the programmes, which accounts for the lower number of respondents ($n = 4$). It's clear that some students are anxious about its use and don't feel that the inclusion of AI activities (at least in their current form) have helped prepare them for their careers, so the risks and affordances of this technology in mathematics programmes will need to be carefully considered. However, there is significant potential in graduate skill development, as illustrated in the manim example discussed in Section 2.2: here generative AI can support mathematics graduates in education and communication roles to produce high-quality materials without the need for extensive, specific technical training that is otherwise beyond the scope of their curriculum. Including some examples of the appropriate use of generative AI in course design has the potential to produce technological agile graduates who can rapidly learn new skills.

A small number of respondents seemed to have a negative view of the reflective activities. Villarroel et al. (2018) argue that "students need to be exposed to a variety of [formative] tasks with diverse performance requirements" to develop evaluative judgement, which is perhaps in tension with giving students choice in assessment. Some students, perhaps those with less confidence, may choose a smaller range of familiar assessment formats and question types, giving themselves less opportunity to develop judgement through reflection. A whole programme approach could therefore be necessary to ensure that for any given assessment students have agency, but also experience a variety of tasks and performance requirements across their whole programme.

6. References

Bransford, J.D., and Stein, B.S., 1993. *The IDEAL problem solver: guide for improving thinking, learning and creativity.* W.H. Freeman & Co.

Cahill, M.J., Hynes, K.M., Trousil, R., Brooks, L.A., McDaniel, M.A., Repice, M., Zhao, J., and Frey, R.F., 2014. Multiyear, multi-instructor evaluation of a large-class interactive-engagement curriculum. *Physical Review Physics Education Research,* 10(2). https://doi.org/10.1103/PhysRevSTPER.10.020101

Carifio, J., and Perla, R., 2008. Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, 42(12), pp.1150-1152. https://doi.org/10.1111/j.1365-2923.2008.03172.x

Code, W., Merchant S., Maciejewski, W., Thomas, M., and Lo, J., 2016. The Mathematics Attitudes and Perceptions Survey: an instrument to assess expert-like views and dispositions among undergraduate mathematics students. *International Journal of Mathematical Education in Science and Technology*, 47(6), pp.917-937. https://doi.org/10.1080/0020739X.2015.1133854

Dell'Acqua, F., McFowland, E. III, Mollick, E., Lifshitz-Assaf, H., Kellogg, K.C., Rajendran, S., Krayer, L., Candelon, F., Lakhani, K.R., 2023. Navigating the jagged technological frontier: field experimental evidence of the effects of AI on knowledge worker productivity and quality. *Havard Business School Working Paper* 24-013.

Demos 2023. *The AI Generation. How universities can prepare students for the changing world.* Available at https://demos.co.uk/wp-content/uploads/2023/11/The-AI-Generation-2.pdf [Accessed 8 February 2025].

French, S., Dickerson, A., and Mulder, R.A., 2023. A review of the benefits and drawbacks of high-stakes final examinations in higher education. *Higher Education*, 88, pp. 893-918. https://doi.org/10.1007/s10734-023-01148-z

Freudenthal, H. 1968. Why to teach mathematics so as to be useful. *Educational Studies in Mathematics*, 1, pp. 3-8. https://doi.org/10.1007/BF00426224

Griffiths, W. and Keith, L. (2021). Communities and narratives in neglected spaces: voices from SMASHfestUK. *Journal of Science Communication,* 20(01), C04. https://doi.org/10.22323/2.20010304

Guilkers, J.T.M., Bastiaens, T.J., and Kirschner, P.A., 2004. A five-dimensional framework for authentic assessment. *Educational Technology Research and Development,* 52, pp. 67-86. https://doi.org/10.1007/BF02504676

Jeong, H.J., and Lee, W.C., 2016. The level of collapse we are allowed: comparison of different response scales in safety attitudes questionnaire. *Biom Biostat Int J.,* 4(4), pp. 128-134. https://doi.org/10.15406/bbij.2016.04.00100

Jones, M., and Megeney A., 2019. Thematic problem solving: a case study on an approach to teaching problem solving in undergraduate mathematics. *MSOR Connections*, 17(2), pp. 54-59. https://doi.org/10.21100/msor.v17i2.978

Jones, M., Megeney, A., and Sharples, N. 2022. Engaging with maths online – teaching mathematics collaboratively and inclusively through a pandemic and beyond. *MSOR Connections,* 20(1), pp. 74-83. https://dx.doi.org/10.21100/msor.v20i1.1322

Jones, M.M., Kazim Ali, Z., Lawrence, S., Masterson, B., Megeney, A., and Sharples, N., 2025. Public engagement for student empowerment. *MSOR Connections,* 23(3), pp. 83-91.

Likert, R. 1932. A technique for the measurement of attitudes. *Archives of psychology*.

Maciejewski, W., Tortora, C., and Bragelman, J. 2021. Beyond skill: students' dispositions towards math. *Journal of Developmental Education*, 44(2), pp. 18-25.

Mason, J., Burton, L., and Stacey, K., 2010. *Thinking Mathematically*. Second edition. London: Prentice Hall.

Masterson, B., Jones, M.M., Megeney, A., Sharples, N., and Lawrence, S., 2023. Authentic by design: developing mathematicians for the talent economy. *MSOR Connections,* 21(2), pp.44-51. https://dx.doi.org/10.21100/msor.v21i2.1400

Masterson, B., Sharples, N., Jones, M.M., Lawrence, S., and Megeney, A., 2024. Authenticity in learning, teaching and assessment. *MSOR Connections,* 22(2), pp.29-36. https://dx.doi.org/10.21100/msor.v22i2.1488

Megeney, A., 2016. 'Engaging with mathematics: how mathematical art, robotics and other activities are used to engage students with university mathematics and promote employability skills', *CETL-MSOR Conference 2015*. University of Greenwich, 8-9 September. Sigma network.

Norman, G. 2010. Likert scales, levels of measurement and the "laws" of statistics. *Adv in Health Sci Educ*, 15, pp.625-632*.* https://doi.org/10.1007/s10459-010-9222-y

Ozimek, D., Good, L., Leggieri, A., Morgante, B., Phillips, M., Watson, G., and Wilk, D., 2024. Exploring prelicensure nursing students' perceptions and attitudes towards mathematics in a concept-based curriculum. *Teaching and Learning in Nursing*, 19(4), pp.e617-e623. https://doi.org/10.1016/j.teln.2024.05.004

Polya, G. (1990). *How to solve it.* Second edition. London: Penguin.

Pretorius, L, van Mourik, G.P., and Barratt, C., 2017. Student choice and higher-order thinking: using a novel flexible assessment regime combined with critical thinking activities to encourage the development of higher order thinking. *International Journal of Teaching and Learning in Higher Education* 29(2), pp 389-401.

QAA, 2023. *Subject benchmark statement: Mathematics, Statistics and Operational Research*. Available at https://www.qaa.ac.uk/docs/qaa/sbs/sbs-mathematics-statistics-and-operational-research-23.pdf. [Accessed 8 February 2025].

Russell Group, 2023. *Principles on the use of generative AI tools in education*. Available at https://russellgroup.ac.uk/media/6137/rg_ai_principles-final.pdf. [Accessed 8 February 2025].

Sanderson, G. 2024. Large Language Models explained briefly. *Youtube.* Available at https://www.youtube.com/watch?v=LPZh9BOjkQs. [Accessed 8 February 2025].

Sharples, N. 2025. MAPSanalysis. *Github.* Available at https://github.com/nicholassharples/MAPSanalysis. [Accessed 14 February 2025].

Sullivan, G.M., and Artino, A.R. Jr., 2013. Analyzing and interpreting data from Likert-type scales. *J Grad Med Educ.,* 5(4), pp.541-542. https://doi.org/10.4300/JGME-5-4-18

Tao, T. 2024. [Mathstodon] 13 September. Available at https://mathstodon.xyz/@tao/113132502735585408. [Accessed 8 February 2025].

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