On the use of technology in university mathematics teaching and assessment in STEM degree schemes.

Sofya Lyakhova

A discussion paper for JMC, June 2023

The paper focuses on the examples of the use of technology in mathematics teaching in HEIs in the UK and is written to support the work of the JMC *Digital Technologies and Mathematics Education Working Group*.

Introduction

Research on university mathematics teaching paints a picture of innovative and diverse use of technologies (see, e.g., Jordan, 2013; Sangwin, 2013; Heeren & Jeuring, 2014; Barana *et al*., 2015; Paiva *et al*., 2017; Graham, 2020; Divjak *et al*., 2022; Hilliam & Calvert, 2019; Kado, 2022; Ozgur *et al*., 2017) with many new research findings emerging during and because of Covid-19 (Bakker et al., 2021; Fitzmaurice & Fhloinn, 2021; Sangwin & Bickerton, 2021). Unlike schools, universities are not bound by high-stakes tests set by external exam boards but are research-driven environments adaptive to both external and internal challenges. These often include attracting suitable applicants but also improving learner experience and facilitating learner progression to employment. Additionally, universities are close to the frontiers of blue-sky research in AI and machine learning which is widely recognised as one driving force behind automated assessment (Hwang & Tu, 2021; Zhai *et al*., 2021). Thus, arguably, universities have reasons but also opportunities to engage more readily in a small-scale innovation in their day-to-day teaching.

This served as a motivation to consider examples of technology implementation in universities with a view that HEI experience may be useful, in a broad sense, when undertaking implementation of technology at any stage of education. The paper presents case studies of technology used in mathematics teaching in four universities in the UK.

Sampling

British universities vary in their approach to curriculum and assessment. Therefore, in our sample we aimed to combine mathematics departments known for their innovation as well as departments that may be perceived as teaching in a more traditional style. The four case studies presented include both modern teaching focused universities and traditional research universities covering a diverse range of student cohorts and degree courses varying from pure mathematics to applied mathematics and statistics. Each case study tells a story of one undergraduate module. The mathematics modules considered are ranging from those that traditionally perceived as pure 'pen and paper' mathematics, to applied modules using specific purpose-built technologies, to modules utilising widely available versatile digital coding tools. Two modules although taught by mathematics departments serve students studying broad ranges of STEM subjects. Three of the cases studies (Geometry, Python and Statistics) focus on both doing and assessing mathematics with technology, and one case study (Linear Algebra) is an example of technology used entirely for assessment.

Method

The study was conducted as interviews with academics leading or co-leading each university module considered. Academics were asked to describe the module and the technology used, explain the rationale of the module and learning outcomes, how the technology influenced or was influenced by the aims of the module. They were also asked about their view of how the use of technology has influenced HEI teaching and learning in the past and how they see it in future, and what changes they wish were made to school education from the point of view of transition to university and in view of wider aims of education. The cases were written by the interviewer paraphrasing what was discussed in the interviews with some original expressions preserved. These were then checked with the academics for factual correctness.

Case studies overview

Despite the modules and universities considered being diverse, some common themes appeared in the interviews:

- Already a decade ago in the universities considered around 30-40% modules constituting a typical mathematics degree, had technology embedded in them. The academics interviewed believe this is not untypical across other universities in the UK. One university participating in the case studies is currently re-structuring their syllabus to take this up to 80%. In another university automated assessment is integrated in all mathematics STEM+ degrees underpinned by a strategic decision taken four years ago to provide a free loan of a laptop to every student.
- Assessment of the modules where technology is taught substantially diverts from traditionally hand-written exams (and those elements of assessment that are perceived as encouraging memorisation) and may include coursework, project, presentations, automated quizzes, electronic workbooks and other electronic as well as hand-written outputs.
- The use of technology is motivated by technology being the main medium through which mathematics is applied in the workplace, but also by the benefits to learners. In particular, in the pure mathematics modules in our sample (Geometry and Linear Algebra) technology was perceived as an effective framework for rethinking how deductive mathematical arguments are constructed and how these are best taught, see also Sangwin & Bickerton (2021) and references within.
- There are explicit expectations of more technology learning at school in the context of doing mathematics both for the benefit of transition to university but also for reflecting the modern nature of mathematics as a discipline as well as wider societal benefits.

Case study 1. Automated assessment in a first-year Linear Algebra module.

Module. ‘Introduction to Linear Algebra’ is a compulsory module for Mathematics students and is available to anyone studying Science or Engineering. The course is typically taken by 800+ undergraduate students.

Vision. Linear Algebra is one of only two courses mentioned explicitly in the QAA benchmark statement and is normally one of the first abstract mathematics course that students meet at university. While students learn to solve systems of linear equations they should also learn about the nature of the discipline and the central role of proof in mathematics. Since the first release in 2005, the automated assessment system that was originally intended as a tool to give timely formative feedback to large numbers of students, has expanded and is now also used for summative assessments. Recently a strategic decision was made to offer the loan of a laptop to any student joining mathematics degrees to ensure 100% of learners have 24/7 access to the technology needed for their study.

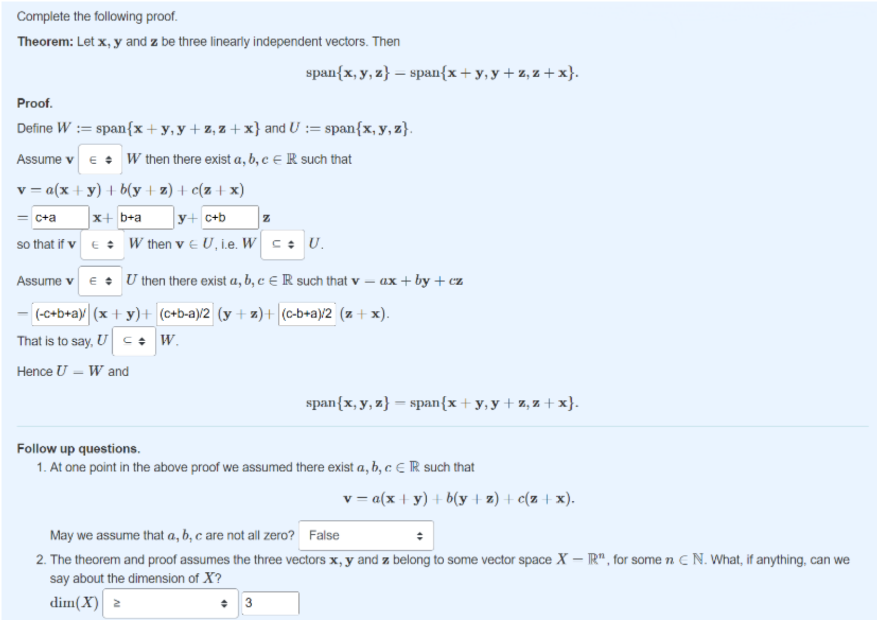
Module content. The module is an introduction to linear algebra, mainly in real vector spaces but concludes with an introduction to abstract vector spaces. The principal topics are vectors, systems of linear equations, matrices, eigenvalues and eigenvectors and orthogonality. The important notions of linear independence, span and bases are introduced. This course is both a preparation for the practical use of vectors, matrices and systems of equations and also lays the groundwork for a more abstract, pure-mathematical treatment of vector spaces.

Learning outcomes include being able to solve systems of linear equations and demonstrate an understanding of the nature of the solutions; perform accurate and efficient calculations with vectors, matrices, eigenvalues and eigenvectors in arbitrary dimensions; demonstrate a geometrical understanding of vectors and vector operations in 2 and 3 dimensions; demonstrate an understanding of orthogonality and projection in arbitrary dimensions; argue in a formal style (definition/theorem/proof or use examples) about statements in linear algebra, as the first step towards a more abstract, pure-mathematical treatment of vector spaces.

The summative assessment consists of continuous weekly assessment quizzes (40%), written hand-in assessment (30%) and a two-hour-time-limited online test (30%). Weekly online assessments are timed (2 hours) and allow one attempt. Best 8 out of 10 counts towards 40% of the course. There are four written hand-in assessments and 3 out of 4 count toward 30% of the course. These may use materials from any of the previous weeks to properly reflect the cumulative nature of the subject. The end-of-module online test includes random auto-graded questions and manual graded (human-marked) responses. There are also reading quizzes and weekly skills practice quizzes where repeated attempts are permitted, with feedback and worked solutions. These formative assessment tests do not directly contribute a grade to the course, but students must score 80% in the practice tests to unlock the corresponding weekly online assessment. The quizzes are incorporated into a short 5–7-minute video recorded by the lecturer where new material and worked examples are introduced.

What technology is used and how the course is taught. Every week students attend two live face-to-face lectures, and a third lecture is replaced by a series of videos and quizzes. A typical pattern of the online quiz content will be a definition followed by a short discussion and video clip of a worked example by the lecturer. Then there will be a quiz inviting the student to check their own understanding of the material. The questions are developed using STACK (System for Teaching and Assessment using Computer Algebra Kernel, https://stack-assessment.org/) which is an open-source online assessment system for mathematics (but also STEM). STACK questions are not limited to multiple-choice responses and question responses can include the input of algebraic expressions. So that students are not penalised for poor programming skills their answers are validated before they are marked. Also students are given feedback that refers to their specific answer and mistake, as if marked by hand. STACK randomises questions, so different students see different variants of a quiz. This type of questions amounts to 60% of all questions used in the Linear Algebra course alongside *Match, True/false* and *Multiple choice* question types. Other types of questions available in the Moodle system include *gapsellect* used, in particular, in courses where proof comprehension is relevant, such as Calculus, proof and problem solving and operational research, and *coderunner* is used, for example, in engineering mathematics, quantitative skills for biologists and programming modules, see Table 1 and Figure 1. Specifically for practicing constructing mathematical justifications, arguments and proof, STACK now offers *faded worked examples*, *explicit assessment of separate concerns* and *comparative judgment* tools, see, e.g., O'Hagan *et al*. (2022), Sangwin & Bickerton (2022) and Sangwin & Kinnear (2021).

Assessing comprehension and proof is challenging as it requires a subjective opinion about whether learners understand the method, or if they simply follow the method to get the right answer. Thus, in addition to true/false, multiple-choice questions and those that accept a typed mathematical expression, essay-type questions are set as part of the weekly assessment tests where students need to hand-write their explanation or a proof, take a photograph of their writing and upload it with the test. The essay questions are human-marked and a subjective decision is made if the proof is correct and complete, the right level of detail is provided and grammar and spelling are correct.


Complete the following proof.

Theorem: Let $\mathbf{x}, \mathbf{y}$ and $\mathbf{z}$ be three linearly independent vectors. Then

$$\mathrm{span}\{\mathbf{x}, \mathbf{y}, \mathbf{z}\} = \mathrm{span}\{\mathbf{x}+\mathbf{y}, \mathbf{y}+\mathbf{z}, \mathbf{z}+\mathbf{x}\}.$$

Proof.

Define $W := \mathrm{span}\{\mathbf{x}+\mathbf{y}, \mathbf{y}+\mathbf{z}, \mathbf{z}+\mathbf{x}\}$ and $U := \mathrm{span}\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$.

Assume $\mathbf{v}$ [∈] W then there exist $a, b, c \in \mathbb{R}$ such that

$\mathbf{v} = a(\mathbf{x}+\mathbf{y}) + b(\mathbf{y}+\mathbf{z}) + c(\mathbf{z}+\mathbf{x})$

$=$ [c+a] $\mathbf{x}+$ [b+a] $\mathbf{y}+$ [c+b] $\mathbf{z}$

so that if $\mathbf{v}$ [∈] W then $\mathbf{v} \in U$, i.e. W [⊂] U.

Assume $\mathbf{v}$ [∈] U then there exist $a, b, c \in \mathbb{R}$ such that $\mathbf{v} = a\mathbf{x} + b\mathbf{y} + c\mathbf{z}$

$=$ [(-c+b+a)/] $(\mathbf{x}+\mathbf{y})+$ [(c+b-a)/2] $(\mathbf{y}+\mathbf{z})+$ [(c-b+a)/2] $(\mathbf{z}+\mathbf{x})$.

That is to say, U [⊂] W.

Hence $U = W$ and

$$\mathrm{span}\{\mathbf{x}, \mathbf{y}, \mathbf{z}\} = \mathrm{span}\{\mathbf{x}+\mathbf{y}, \mathbf{y}+\mathbf{z}, \mathbf{z}+\mathbf{x}\}.$$

Follow up questions.

1. At one point in the above proof we assumed there exist $a, b, c \in \mathbb{R}$ such that

$$\mathbf{v} = a(\mathbf{x}+\mathbf{y}) + b(\mathbf{y}+\mathbf{z}) + c(\mathbf{z}+\mathbf{x}).$$

May we assume that a, b, c are not all zero? [False]

2. The theorem and proof assumes the three vectors $\mathbf{x}, \mathbf{y}$ and $\mathbf{z}$ belong to some vector space $X = \mathbb{R}^n$, for some $n \in \mathbb{N}$. What, if anything, can we say about the dimension of X?

$\dim(X)$ [≥] [3]

Figure 1. An example STACK question asking students to complete a proof with gaps.

In summary, a pragmatic approach is taken to assessment and while largely skill-based questions are assessed automatically, human-marking is preserved for assessing mathematical justifications and explanations which are submitted as traditional hand-written work.

Question type	Description and key features	Examples of undergraduate modules where used
STACK	The key features include randomly generated variants of questions, input types being an algebraic expression, multiple choice and scientific units and careful validation of student answers, to prevent penalising a student for syntax errors. STACK questions are multi-part with an answer typically provided in the form of a mathematical expression. STACK is able to establish mathematical properties, e.g., equivalence with a correct answer and outcomes include numerical marks as well as a formative feedback.	Used in nearly 90% (32 out of 36) of the modules in one HEI ranging from Pure and Applied Mathematics to Propagability and Statistics to Engineering Mathematics and Biology.
Match	Enables assessment of short free text answers, such as in a single sentence (Jordan, 2012). Teacher could pre-specify an 'allow' and 'deny' list. A student answer is deemed to be correct if a closest match is in the 'allow' list and withing prescribed threshold 'distance' of a string in the allow list. For example, depending on the context a teacher could set '0.5' and '½' to be in the allow list but 'a half' to be in the deny list. Examples of university modules using *match* type of questions in online assessment are algebra and calculus, probability, pure mathematics, combinatorics and graph theory and courses for natural sciences and engineering among many others.	Algebra and Calculus; Probability and Statistics; Proofs and Problem solving; Fundamentals of Pure Mathematics; Symmetry and Geometry; Operational research Combinatorics and graph theory; courses for natural sciences and engineering.
Gapselect	Refers to purposely omitting particular algebraic expressions in a proof for the purpose of asking students to provide that information by filling in gaps. Example of courses using this question type include engineering mathematics, statistics, calculus, quantitative data analysis, mathematical biology and linear algebra.	Engineering mathematics, Statistics, Calculus, Quantitative data analysis, Mathematical biology; Linear algebra.
CodeRunner	Assessment of short fragments of programme code. Although initially referring to computer programming courses, CodeRunner can be used more widely to grade any questions answered with text and for which suitable assessment criteria is developed.	Computing and numerics; Programming; Engineering mathematics; Qualitative Skills for Biologists; Python Programming.

Table 1. Types of online Moodle integrated e-assessment questions and examples of undergraduate modules (see, e.g., O'Hagan et al., 2022 and Sangwin & Bickerton, 2022)

How the use of technology is driven by the aims of the module. Learning to construct a proof is one of the main aims of the module and the one which is considered challenging to automate. Therefore '*faded worked examples*' with blanks for students to fill in the missing parts of the arguments are used in quizzes to teach the proof. While initially, these reduce cognitive load (and hence increase cognition), ultimately students need to become completely independent and to be able to compose the whole proof, including writing the structure, themselves. Therefore, a sequence of 'faded worked examples' with support structure progressively diminishing has been developed.

Teaching challenge and innovation. Automated assessment must be written with the limitations and affordances of the tool in mind, so a profound understanding of the assessment tools as well as the subject matter is essential. Research about proof comprehension as part of machine learning is advancing and well-designed teaching and learning tools could be very powerful. Ongoing accompanying pedagogical research helps to improve tools further as answers are available retrospectively for analysis. For example, an 80% score on practice tests was set in response to how misconceptions exhibited in the assessed quizzes correlated with misconceptions exhibited in the practice tests. Security of assessment, e.g., avoiding plagiarism and impersonation, is important and remains to be best guaranteed by conducting tests (written or online) in traditional invigilated settings.

Student perception. Learners cope well with the automated assessment system. The elements of the flipped classroom are appreciated for the flexibility it offers to learners (e.g., pre-recorded video and automated tests that can be accessed anytime and anywhere), while regular face-to-face lectures offer structure to help students who appreciate a more guided approach to study (e.g., knowing when to sit down and do the work).

Opportunities for school education. It is hoped that the benefits of automated mathematics assessment can be utilised in pre-university education in the next 5-10 years. The flexibility of taking an automated GCSE mathematics test (or a series of tests) when each individual student is ready, as an alternative to a test sat at a fixed prescribed age and fixed prescribed time of the year, would provide much more flexibility. Flexibility in taking tests online when ready, rather than annually or biannually, could be an important step towards achieving a societal goal of everyone being confident using mathematics and numeracy throughout their lives.

Case study 2. Teaching Introductory Statistics module with Minitab statistical software.

Module. 'Introduction to Statistics' is taught to all mathematics and economics students but also on degree courses where statistics is essential, such as business, engineering, and natural and social sciences. Among the 2000+ students that take the module every year, there are diverse groups including those with disabilities and limited access to technology.

Vision. Alongside with teaching statistical techniques, the primary aim of the module is teaching statistical literacy, e.g., equipping students to think about data collection, reading, and communicating statistics and questioning the validity of what is communicated in various sources such as everyday news. The software package used needed to display all the typical features of the technology used in the world of work and career, yet be affordable and easy to teach, without imposing a steep learning curve. The preference was for learners to concentrate on what they are doing with the data, such as data collection via surveys and sampling, rather than learning ins and outs of a sophisticated analysis. Given the large and diverse cohort of students the choice of technology also aimed not to disadvantage less computer enthusiastic students, and equivalent forms of assessment for learners with limited access to technology had to be put in place. The module was first introduced in its current from in 2013.

Module content is split into five sections which are descriptive statistics, regression and surveys, hypothesis testing, association and estimation, experiments and clinical trials. There are online resources to help support students' study, including screencasts (short audio-visual presentations by the lecturers), practice quizzes, interactive computer resources and a computer book.

What technology is used. Minitab (www.minitab.com) statistical software is used in the module to carry out some statistical computations and produce statistical analysis. Minitab is regarded to be a leading statistical education software but is also used as a statistical analysis tool in some industries.

Learning outcomes include knowledge and understanding of key ideas on statistics, basic statistical vocabulary, and a repertoire of statistical techniques for analysing data. Students learn to select and use appropriate techniques and strategies for analysing data in a range of everyday contexts, to interpret statistics in real-life situations, to provide answers in a non-technical format and to develop simple statistical arguments. In term of practical skills, learners are expected to describe questions about data statistically, to use a computer to analyse data, to analyse and comment on statistical analyses. Students should be able to explain statistical ideas in writing, use appropriate terminology, notation and style, but also to develop skills in learning independently, e.g., managing study times, learning actively, reflecting on progress and planning further learning, and to use ICT tools, such as the electronic assessment system and online resources.

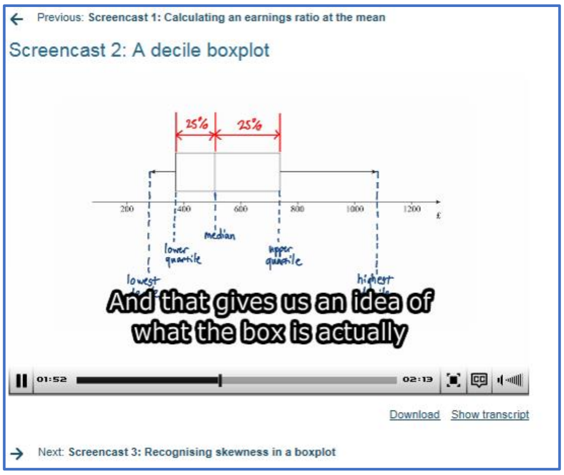


Figure 2. A screenshot of a screencast explaining boxplots

The summative assessment has four tutor-marked assignments (38.75%), three interactive computer-marked assignments (26.25%) and the end-of-module assignment (35%). Computer-marked assignments are similar to the practice quizzes that students can access online and are linked to each unit of study within the module, and also play a role in providing formative feedback. Learners can change their answers as many times as they like until they submit the assignment. The score for each computer-marked assignment is displayed usually about a month after the submission date for that assignment. Each student is given different questions as part of the interactive computer-marked assignments, but all students are given questions of the same level of difficulty. Learners receive automated feedback, in addition to a grade. The end-of-module assignment relates to the whole module and consists of two parts, one final computer-marked assignment and one written assignment. Formative feedback is given as part of students attempting practice quizzes and interactive computer resources.

How the use of technology is driven by the aims of the module. Technology is considered essential for learning statistics which is reflected in the learning outcomes. Therefore, the assessment is designed in such a way that learners necessarily engage with Minitab and other technology. However, in exceptional circumstances some data could be made available in alternative formats to such an extent as to help students to pass the module. By design either equivalent forms of content and assessment are available elsewhere, or a particular technology-heavy bit is not vital to get through the course. Various adjustments are made for students with disabilities, such as screen readers for visual impairment.

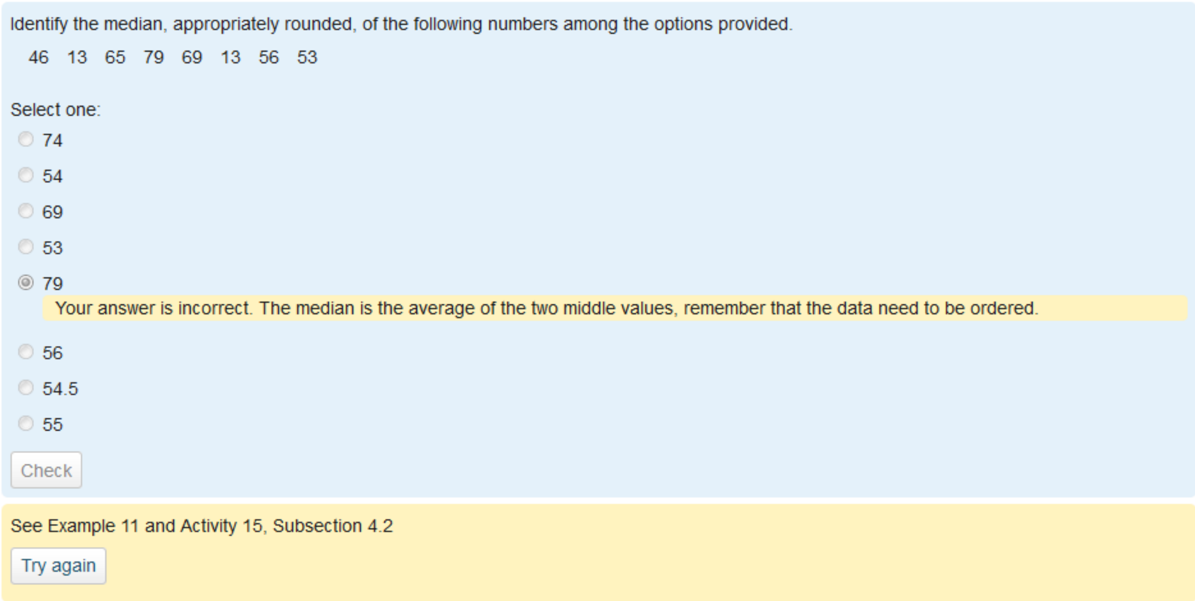


Figure 3. An attempt at a practice quiz question (image cropped to show answers and feedback only)

Student perception. The module has one of the best pass rates of the level one courses despite being taught across a large spectrum of degree courses. This is attributed partially to its overall aim and engaging materials. In terms of technology, there are always students who dislike technology, and these are assigned a tutor for one-to-one support to overcome this. Students have a choice of modules to study more or less technology as part of their degree, but on average a third of a mathematics degree modules incorporate doing mathematics with technology.

Teaching challenge and future innovation. Recently there has been a big shift in statistics and in data science, with one professional package being often used to perform all statistical work from start to end and alongside an online notebook. The degree courses therefore need to reflect this. An ongoing rolling process of rewriting the statistics syllabus will gradually result in an increasing amount of technology use and in accommodating larger and more complicated data sets. When choosing statistical data packages, balancing affordability, ease of use and powerful qualities against transferable skills and limited time learners spend studying for a degree, requires careful consideration.

Hopes for school education. Ultimately, teaching statistics with pen and paper should give way to playing with datasets. Upskilling teachers for using various statistical software may be difficult to afford in the short term, but more emphasis on using Excel, or similar widely available technology, should be considered. Excel is excellent for teaching skills useful in any job, such as graphs and manipulating data. Given the importance of statistical literacy, more thought should be given to developing free statistics software packages appropriate for school use in future.

Case study 3. Teaching Python-based programming to first-year mathematics undergraduate students

Module. ‘Programming’ is one half of the ‘Programming, Foundations and Connections’ compulsory module for Mathematics, Mathematical Sciences and Mathematics and Statistics undergraduate degree programmes studied in the first year.

Vision. The predecessor to this module was developed ten years ago to overcome the challenge of students not knowing how to program when entering mathematics degrees. While programming was seen as crucial for some applied modules in the 2nd and 3rd year, the Programming, Foundations and Connections module was developed to teach first-year students (object-oriented) programming in the context of applied and pure mathematics. Students learn not only how to write computer code but to critically think about algorithms, computations, and complexity and all in the context of doing mathematics. The overall aim is for students to become confident in implementing and understanding mathematical algorithms. Technology-related modules constitute approximately 40% of what is learnt in the first year and – dependent on the student specialisation - across the whole degree course.

Module content. Mathematical content is confined to simple numerical problems with integer numbers, such as writing a code to find a factorial of a number. The programming syllabus covers programming paradigms, building elements (variables, data types, arrays and strings), control flow (e.g., loops, functions and recursion), understanding and analysing algorithms (e.g., complexity analysis, divide-and-conquer methods, big-O notation), sustainable software engineering and working with data and libraries.

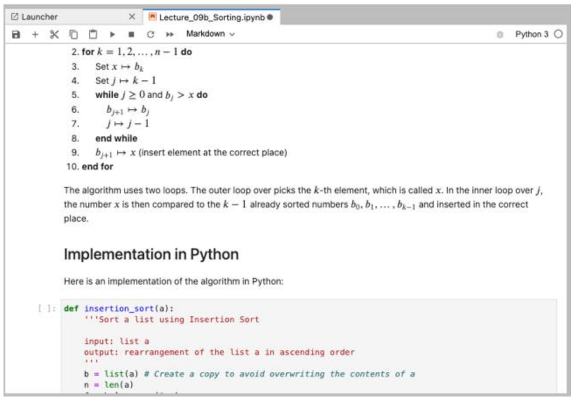


Figure 4. A screenshot of the Jupyter notebook.

Learning outcomes include students writing pseudo-code and implementing algorithms in Python; applying modern programming paradigms to mathematical problems; demonstrating an understanding of principles of software design; analysing the complexity of algorithms; and manipulating and analysing data.

The summative assessment consists of two coursework assignments, each worth 25% and each being a set of programming exercises. Students are given an algorithm that is written on a piece of paper and have to implement it, that is, to write a Python code for it. Then the algorithm must be applied to sets of inputs, for example, a sorting algorithm is applied to sort a given set of numbers. In addition, students are given a small set of simplified tests that they can use to check their implementation before submission. Students submit a code (in a form of an electronic iPython notebook) together with any comments they write about the code. Marking the code is done semi-automatically, and each code is run through further automated tests in addition to manual checks for coding style (e.g. comments and formatting). Formative assessment consists of weekly programming exercises, for which the students receive feedback in the weekly computer labs. Again, simplified tests are available to students so they can run their codes and get immediate feedback as well.

How the use of technology is driven by the aims of the module. Mathematics is well-integrated with programming so that the main elements of object-oriented programming can be learnt in the context of mathematics. For example, the mathematical relationship between fields and rings can be expressed in terms of a Python class hierarchy. The module initially incorporated MATLAB which is a programming tool commonly used by engineers and by applied mathematicians and physicists. Two years ago, it was decided to switch to Python which by then became one of the most popular programming tools in STEM. Typically taught in Computer Science, Physics and Engineering degrees courses, the use of Python is gradually becoming more popular in mathematics. There are examples of Python used in mathematics but also in economics degree courses by leading universities and more are expected in future.

Student perception of the module is very positive although students' programming skills differ and while some find learning programming easy, there are students who find it challenging. Students who took Python in school have an advantage. The skills acquired are appreciated by the students. These are perceived as essential for the modules studied later, for an industrial placement in the third year (in the past, students have been asked to do a presentation to the 1st year undergraduates about the placement and its connection to the programming module) and are seen as generally important for their future career. When switching from MATLAB to Python, the student preference was for the latter by a large margin because Python is widely used (including being taught in school) and is perceived as giving students transferable skills.

Teaching challenge. Switching to Python required no change in mathematics but all codes needed to be rewritten and student access to the Python programming environment to be provided. The latter was developed using iPython notebooks, which are run on a dedicated Jupyter server. Setting up this server and making sure that it scaled to large numbers of students required a lot of technical work behind the scenes. One advantage of this approach is that students can access the server through a web browser from any device without the need to install specialist software. For the unit convenor, this monolithic programming platform simplifies the distribution of teaching material to the students.

Independent of the particular changes to this module, it was further decided to restructure the whole curriculum and offer an additional degree programme that is focused on data science and machine learning where most of the modules will incorporate code-writing. It is envisaged that 80% of these modules may be Python-based in the future.

Hopes for school education/transition from school to university. It is hoped that as Python could be learnt in school, more students will be entering degree programmes with programming skills. This would allow for rapid learning of applying programming specifically to mathematics.

Case study 4. Teaching the structure of mathematics and mathematical communication with mathematical typesetting software in the context of a first year Geometry, Logic and Communication module.

Module. 'Geometry: Mathematics, Logic and Communication' is a compulsory first year module taught in the first term.

Vision. The module aspires to challenge students' perceptions about mathematics. Mathematics is much more than a collection of techniques for solving equations. It is a system for organising our thoughts, reasoning, describing and analysing the world around us, and communicating precise conclusions. Mathematics is fundamentally about constructing an argument, so it can be shared with others (such as in the form of a typed text in a book or online) and can persuade somebody else about its statements. Hence, in learning to communicate mathematics, there should be an emphasis on all types of communicating mathematics, hand-written, typed and oral, synchronous (live presentation) and asynchronous (peer assessment and feedback after reading a typed text). Historically, programming and communicating skills were taught as part of a generic employability module, while students were exposed to the structure of mathematics indirectly through various mathematics modules. A module on comparatively non-advanced geometry where all elements of mathematical structure could be showcased, represented an opportunity to combine all of the above and introduce students to programming in the first year of their degree.

Module content. Students start with the familiar geometry of the classical Greek period before being introduced to more modern ideas such as non-Euclidean geometry, cartesian coordinates, and the real number line. Learners practice written and oral mathematical communication through the use of the LaTeX typesetting system and a drawing tool called TikZ. Importantly, students are expected to write mathematics in sentences and as a piece of coherent text as one may find mathematics in textbooks, research literature or scientific reports.

Learning outcomes include students solving problems in Euclidean geometry and other geometries, constructing basic logical proofs using various axiom systems, typesetting mathematics using LaTeX, including diagrams where appropriate, giving oral presentations on their work, demonstrating an understanding of the differences between well-written mathematics and poorly-written mathematics.

The summative assessment has four components - coursework (10%), presentation (10%), a mathematical writing project (20%) and an open-book exam (60%). Technology is assessed in all parts. Formative assessment is a mixture of peer assessment and lecturer assessment. As part of peer assessment students submit their original work and a summary of comments received from peers together with their views on how to improve using the comments. All the components of the summative assessment are human-marked.

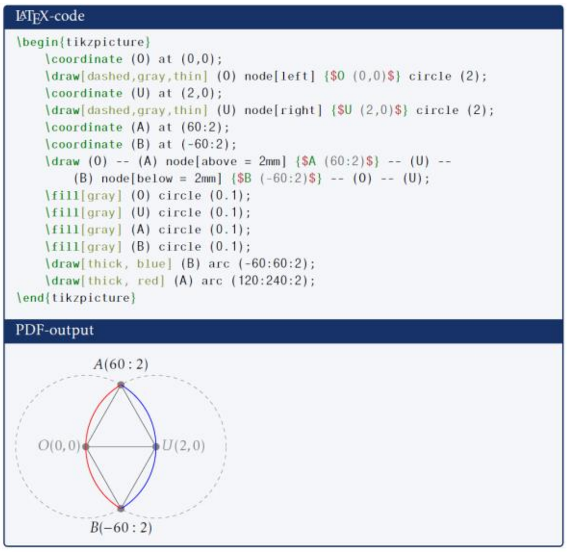


Figure 5. An example of LaTex input and PDF output

What technology is used how its use is taught. The emphasis in first few weeks is on learning mathematical language and structure, e.g., the place of axioms, definitions, theorems, corollaries, proofs in a mathematical text. This is done in the context of low-level geometry that students are familiar with from school. Simultaneously, students learn how to structure a piece of mathematical text using LaTeX, which is a technical high-quality typesetting system designed to produce scientific pdf-type documentation using mathematical symbols and formulae. While moving to new unfamiliar branches of geometry, students also learn to programme mathematical symbols and formulas in LaTeX and later, to incorporate drawings using TikZ, a tool to create graphic elements in LaTeX. TikZ is used for drawing lines, dots, curves, circles and other elements used for visualising mathematical concepts and requires higher proficiencies from learners than typing sentences and formulae in LaTeX . Towards the end of the course, students are expected to solve problems, perhaps, initially using pen and paper, then structure their argument in accordance with the convention of mathematical writing and design diagrams to illustrate their argument, and, finally, to typeset the argument in LaTeX and present it to peers and teachers as a neat, clear and easy to follow piece of text. To emphasise that mathematics is written for the purpose of others understating it, peer assessment including giving and getting feedback, is used throughout the module.

How the use of technology is driven by the aims of the module. While in theory it is possible to teach the structure of mathematics with no technology, LaTeX provides a ‘natural’ framework to review and structure one’s argument, such as when choosing the correct set of commands (called ‘environments’) to display a definition, an axiom, or a proof. It also encourages clearer presentation of mathematics which is better suited for peer assessment than hand-written notes. Additionally, enabling students to type a neat portfolio of their work that they could proudly present to future employers is envisaged to contribute to employability skills. GeoGebra, a free interactive geometry, algebra, statistics and calculus application, was considered as an alternative to TikZ as already familiar to students from school, however there are disadvantages when compared with TikZ.

Student perception of the module has evolved since it was first introduced five years ago. Initially, many students resisted the challenge of typing mathematics and saving and submitting their assessments electronically as it contradicted their school habits of doing mathematics. Gradually, there have been more students with programming skills or wanting to learn to program. Currently some students see this module as their favourite allowing them to display their creativity, such as through introducing colour and more aesthetics in the presentation. Typically, once completing the course, there would be some students who choose to continue typing their notes and homework solutions neatly throughout the course of their degree, while others would revert to handwriting mathematics. All students will use LaTeX for typing and presenting a project in the final year of their degree.

Initial teaching challenge. Teaching programming required a different methodology in comparison with teaching mathematics.

Hopes for school education. While approximately a quarter of all modules in our BSc and MSc degree courses incorporate using maths-specific technology, there are students who find programming challenging and who do not perceive technology to be essential for communicating mathematics. It is hoped that, with the most recent school reforms emphasising mathematical communication, there will be a place for learning mathematics-specific technology as part of learning to communicate mathematics in schools in future.

Discussion

We considered four mathematics modules from different and diverse higher education institutions teaching across four jurisdictions in the UK. While we do not wish to claim that our sample of HE mathematics department is typical, it exemplifies prolonged innovation and varied use of technology at university level.

Each year the modules considered in the paper are taken by students studying mathematics, statistics, economics, business, natural science and engineering degree courses. The modules include pure mathematics that could be taught with no technology as well as applied mathematics and statistics modules where technology is essential. In terms of technology, we have Minitab which is a purpose-built statistical software popular in education but less popular in the workplace, but also Python which is a versatile programming tool widely used in workplace. LaTeX is a unique editor tool from early 90s developed for typesetting mathematics and used by professional mathematicians, physicists, computer scientists and engineers but also publishers and recently has been adapted for writing music scores (see, e.g., Suyanto, 2019). TikZ is a graphics package that allows to create high quality diagrams thus allowing to draw geometric shapes when incorporated with LaTeX. The latter on its own does not have a drawing capacity. Finally, the Linear Algebra case study features a computer algebra system integrated with Moodle for assessment.

Technology seems to be well-embedded in the mathematics teaching in all universities considered. Both the Statistics and the Python case highlight the effort but also the skill required to choose the right technology. The significance of switching from MatLab to Python in our case studies should not be underestimated. Although MatLab can be also used as a programming language, it is most commonly utilised as a numerical computing environment and thus switching from MatLab to Python in our case study requires the students to develop new valuable skills. They have to find their own way around a more flexible and powerful but less monolithic programming ecosystem. Ozgur and colleagues (2017) considered using Python, MatLab and R for teaching operations research and statistics to students in a college setting and concluded that Python was the best programming language to be taught in a classroom environment. This was because Python is easy to use and it allows easy student access as it is open source. However, according to Ozgur *et al*. (2017) studying R may offers a competitive advantage when looking for a graduate job, while MatLab may encourage users to learn programming in depth exactly because it is not an open source.

Despite the differences in the mathematical nature of the modules and the digital technologies used, the rationales for teaching doing mathematics with technology are similar and include pragmatic but also societal, e.g., employability, and pedagogical aspects. Technology is perceived as the main medium through which mathematics is applied in the workplace therefore learning it contributes to preparing students for work. But technology is also viewed as an effective (convenient, appropriate, relevant) framework for rethinking teaching and assessment. As we have seen in the Geometry, Linear Algebra and Python case studies, doing mathematics with is thought to encourage logical/mathematical thought and help students understand mathematical constructs.

We observe that assessment of the modules where technology is taught, substantially diverts from traditionally hand-written exams and may be expected to include several components: some hand-written, some coursework, project or presentation. Some elements are human-marked and others are auto assessed. Although many automated assessment systems for mathematics are available, e.g., STACK, NUMBAS, DEWIS, our study only featured STACK. We refer interested readers to an overview by Jordan (2013) on e-assessment where, among other things, STACK is introduced in the context of other computer-marked assessment system. We remark that while some important underpinning research enabling the Moodle types of questions mentioned in the Linear Algebra case study are not new, the research area keeps advancing with some emerging results being due to the Covid-19 changes of practice. Indeed, good reliability of the computer marking of short-answer free-text responses of around a sentence in length has been established more than a decade ago, see, e.g., Butcher & Jordan (2010). CodeRunner types were developed in 2016 (Lobb & Harlow, 2016) and first introduced for teaching in a UK university in 2017. Comparative Judgment is becoming a well-established research tool for data collection and validation (Bisson *et al.*, 2016; Jones *et al*., 2019; Davies *et al*., 2020). STACK was originally designed for formative assessment and was previously found to be unsuitable for automating questions from existing school exams (Sangwin & Kocher, 2016). In 2017 an online mock STACK exam was implemented for re-sits at one university and where it is now used annually. Further during Covid-19 STACK online synoptic assessment has been reported to replace the traditional exams. Most recent research claims that STACK could be successful in high-stakes situations (O'Hagan *et al*., 2022).

In addition to true/false and multiple-choice familiar to most readers, the Linear Algebra case study referred to some other types which we summarised in Table 1. We can see from the table that auto-assessed questions are utilised not only in the fields of mathematics associated with computation, coding and modelling but also in the modules associated with constructing mathematical arguments, and which may be perceived as less suitable for auto-assessment. Indeed, according to one of the interviewed academics, there have been recently interesting developments in adapting automated assessment for teaching mathematical proof and aiding comprehension. Writing an explicit proof in an exam has been associated with memorisation and recently, assessing proof comprehension gained popularity instead. This can be done through asking learners to give a situation when the proof fails, to justify why certain examples do/do not exist, to provide or select examples satisfying certain properties used in the proof and so on. Such questions could be formalised and to some extent assessed with the categories of questions considered above, including multiple choice and true/false questions but some new approaches are also being explored.

For example, STACK offers 'faded worked examples' and 'explicit assessment to separated concerns' type questions as one approach to help learners to practice constructing mathematical arguments. Classic faded worked examples refer to a progressive sequence of worked examples in which steps within a worked example are systematically removed, requiring students to take increasing responsibility for completing the problem. It is considered useful to practice faded worked examples before starting to solve problems independently. A checklist for writing a proof comprehension questions has been added recently to help teachers to construct their own questions. Explicit assessment of separated concerns is defined as explicitly identifying, teaching and assessing specific topics in relative isolation in anticipation of their immanent need, e.g., the sigma notation or P(n+1) statement in the proof by induction.

Comparative judgement is another assessment tool that could aid comprehension. Based on a well-established psychological principle that a comparison of two items is more reliable rather than judgment of a single item in isolation (Thurstone, 1927), comparative judgment allows judges to decide which item of the two presented is better. The tools could be used in a variety of ways, e.g., with markers acting as judges or students acting as judges of their work or their peers' work to inform a class discussion.

Wider questions about assessment include the security of online assessments with issues surrounding plagiarism and impersonation, see, e.g., a joint statement from RSS, LMS and IMA released recently (RSS, 2022). One view, which is supported by the assessment strategies in some of our case studies, is that security remains to be best guaranteed by conducting tests (written or online) in traditional invigilated settings. In the cases considered in this study generally, a pragmatic approach is taken to assessment and while largely skill-based questions are assessed automatically, human-marking is preserved for assessing mathematical justifications and explanations which are submitted as traditional hand-written work.

Implications for school education and concluding remarks

We propose that the findings of our study suggest several reasons for more technology learning in school, but they also highlight considerable challenges associated with implementing it.

One of the reasons for introducing teaching mathematics with technology could be an improved transition to HE. All the interviewed academics wished for more mathematics with technology learning at school. This is especially relevant if universities are to embed more technology in their mathematics teaching in near future, as exemplified in our case study. But one may also argue that the very reasons for HE for introducing doing mathematics with technology, as discussed above, are relevant to the aims of school education. Like universities, schools are preparing students for the world of work and thus need to reflect the ways in which mathematics is used in workplaces. The Statistics case, in particular, highlighted how adapting the module to reflect how statistics is practiced in the real world remains a challenge and that authentic assessment would require the use of technology in schools. Similarly, the fact the use of technology can encourage logical mathematical thinking and help students understand mathematical constructs is relevant at school level as much as it in HE (see also Kopcha *et al*., 2016).

Many studies have argued for more technology in statistics teaching. A suggestion in the present paper to use Excel to fulfil this, could be argued to be beneficial from a different point of view. Ozgur *et al.* (2017) in a sideway discussion about SPSS, SATA, SAS and Excel noted that Excel is probably the most commonplace software used for performing everyday analytics by small businesses. Thus, studying Excel is beneficial for wide groups of students. In relation to making a better use of technology in statistics teaching and learning at school, we remark that the latest A-level qualification reform in England and Wales aimed to address this, but we are yet to see the true impact of the reform. In Scotland a new pilot Higher Applications of Mathematics level 6 qualification was introduced in 2021 and first evidence is expected to emerge in a few years' time.

Numerous challenges could be associated with introducing technology into mathematics teaching. Not only a considerable initial effort but continuous innovation seems to be required.

Joint Mathematical Council of the UK

- The interdisciplinary nature of doing and teaching mathematics with technology has potential implications on teacher professional development in terms of assumed extensive knowledge of the landscape of the discipline but also knowledge of specific technology used. For example, it follows from our findings that if schools are to incorporate technology in the learning of maths/stats they need to be flexible and agile like universities are in changing technologies (such as MatLab to Python) to reflect practice in the world of work.
- In terms of teaching the use of technology (rather than teaching through technology) universities design teaching, technology use and assessment around each other themselves. As we have seen, HEIs change the assessment to drive students to engage with particular aspects of the teaching. Schools do not have that luxury and must ensure students have studied an externally set syllabus to ensure they can pass externally set assessment. One may hypothesis that unless those assessments encourage the use of technology, a school would not invest time in reshaping their curriculum.
- Access to technology needs to be considered. Both the Statistics and Linear Algebra case spotlight the issues of access to technological devices and the corresponding problems of equity.
- The task of revising the mathematics curriculum to include technology is onerous. In this respect one needs to consider which layer of the school education system is better suited to do this given that individual mathematics departments in schools do not have the same level of resource as in universities.

Yet, opportunities for integrating technology into the learning of mathematics seem to exist already. There is evidence of HEIs adapting their courses to include technology that is learned in (at least) some schools, such as Python. Yet, we are not aware of the systematic use of Python for solving mathematics problems at school. The Geometry case serves as a reminder that professionally word processed documents and not scribbled notes of calculations are the main means of communicating mathematics for work, study and leisure. Visualizing mathematical concepts with pictures aids the clarity of what is communicated, and also comprehended. Mathematics communication is featured in the school mathematics curriculum (see e.g., CCEA, 2023; Education Scotland, 2017; Ofqual, 2017; Welsh Government, 2019) and could be explored with dynamic geometry tools already available in school classrooms, such as GeoGebra.

Finally, examples of good use of automated assessment are available, and we propose that some features of automated assessment exemplified in the paper, could be applicable and adaptable at GCSE and post-GCSE stages of school education which ideally needs separate investigation. The proposition prompted in one of the case studies of an automated GCSE mathematics assessment is particularly exciting. A test taken when each individual student is ready, as an alternative to a test sat at a fixed prescribed age and fixed prescribed time of year, would be an important step towards achieving a modern societal goal of everyone being confident using mathematics and numeracy throughout their lives.

References

Bakker, A., Cai, J., & Zenger, L. (2021) Future themes of mathematics education research: an international survey before and during the pandemic. Educational Studies in Mathematics, 107(1), 1-24.

Barana, A., Marchisio, M., & Rabellino, S. (2015) Automated assessment in mathematics. In 2015 IEEE 39th Annual Computer Software and Applications Conference (Vol. 3, pp. 670-671).

Bisson, M. J., Gilmore, C., Inglis, M. & Jones, I. (2016). Measuring conceptual understanding using comparative judgement. International Journal of Research in Undergraduate Mathematics Education, 2(2), 141–164.

Butcher, P. G. & Jordan, S. E. (2010) A comparison of human and computer marking of short free-text student responses. Computers and Education, 55(2), 489–499.

Davies, B., Alcock, L. & Jones, I. (2020) Comparative judgement, proof summaries and proof comprehension. Educational Studies in Mathematics, 105(2), 181–197.

Divjak, B., Žugec, P., & Pažur Aničić, K. (2022) E-assessment in mathematics in higher education: a student perspective. International Journal of Mathematical Education in Science and Technology, 1-23.

Education Scotland. (2017) Curriculum for Excellence: Numeracy and Mathematics. [online] Available at: https://education.gov.scot/media/sz2lnh1g/numeracy-maths-eo.pdf (accessed June 2023).

Fitzmaurice, O., & Ní Fhloinn, E. (2021) Alternative mathematics assessment during university closures due to Covid-19. Irish Educational Studies, 40(2), 187-195.

Graham, C. (2020) Assessment of computing in the mathematics curriculum using Numbas. MSOR Connections, 18(2).

Heeren, B., & Jeuring, J. (2014) Feedback services for stepwise exercises. Science of Computer Programming, 88, 110-129.

Hilliam, R., & Calvert, C. (2019) Interactive statistics for a diverse student population. Open Learning: The Journal of Open, Distance and e-Learning, 34(2), 203-210.

Hwang, G. J., & Tu, Y. F. (2021) Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. Mathematics, 9(6), 584.

Jones, I., Bisson, M., Gilmore, C. & Inglis, M. (2019). Measuring conceptual understanding in randomised controlled trials: Can comparative judgement help? British Educational Research Journal, 45(3), 662–680.

Jordan, S. (2013) E-assessment: Past, present and future. New Directions, 9(1), 87-106.

Kado, K. (2022) A teaching and learning the fundamental of calculus through Python-based coding. International Journal of Didactical Studies, 3(1), 15006.

Kopcha, T. J., Rieber, L. P., & Walker, B. B. (2016) Understanding university faculty perceptions about innovation in teaching and technology. British Journal of Educational Technology, 47(5), 945-957.

Lobb, R. & Harlow, J. (2016) Coderunner: A tool for assessing computer programming skills. ACM Inroads, 7(1), 47–51.

Ofqual. (2017) GCSE Subject Level Conditions and Requirements for Mathematics. [online] Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/592159/GCSE_Subject_Level_Conditions_for_Mathematics_Feb_17.pdf (accessed June 2023).

O'Hagan, S., Sangwin, C. J., & Zerva, K. (2022) A report on the use of the STACK service in Edinburgh, 2021–22. Technical report, University of Edinburgh.

Ozgur, C., Colliau, T., Rogers, G., & Hughes, Z. (2017) MatLab vs. Python vs. R. Journal of data Science, 15(3), 355-371.

Paiva, R. C., Ferreira, M. S., & Frade, M. M. (2017) Intelligent tutorial system based on personalized system of instruction to teach or remind mathematical concepts. Journal of Computer Assisted Learning, 33(4), 370-381.

RSS. (2022) Statement on Methods of Assessment in the Mathematical Sciences. [online] Available at: https://rss.org.uk/RSS/media/File-library/Policy/Higher%20education%20policy/2022/Statement_on_Methods_of_Assessment_in_the_Mathematical_Sciences_21_09_22.pdf (accessed June 2023).

Sangwin, C. (2013) Computer aided assessment of mathematics. OUP Oxford.

Sangwin, C. J., & Bickerton, R. (2021) Practical online assessment of mathematical proof. International Journal of Mathematical Education in Science and Technology.

Sangwin, C. J., & Kinnear, G. (2021) Investigating insight and rigour as separate constructs in mathematical proof. https://doi.org/10.35542/osf.io/egks4

Suyanto, Y. (2019) The Dynamics Symbol Representation Features of Numeric Music Notation in LATEX. In 2019 5th International Conference on Science and Technology (ICST). Vol. 1, pp. 1-6.

Thurstone, L. L. (1927). A law of comparative judgment. Psychological Review, 34(4), 273–286.

Welsh Government. (2019) Area of learning and Experience: Mathematics and Numeracy. [online] Available at: https://hwb.gov.wales/curriculum-for-wales/mathematics-and-numeracy (accessed June 2023).

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., ... & Li, Y. (2021) A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. Complexity, 2021.

CCEA. (2023) Mathematics and Numeracy. [online] Available at https://ccea.org.uk/key-stage-3/curriculum/mathematics-numeracy (accessed June 2023).