Proceedings of

International Meeting of the STACK Community 2024

Amberg, Germany

11.03.2024 - 13.03.2024

Editor: Dr. –Ing. Michael Weinmann, OTH Amberg-Weiden, Kaiser-Wilhelm-Ring 23, 92224 Amberg, Germany

doi: 10.5281/zenodo.12755221

Foreword

"How many participants can we expect?" was probably the first question we asked ourselves during the organization of the International Meeting of the STACK Communnity 2024. Many more were to follow. At least as many decisions had to be made before the three-day conference could start on March 11, 2024.

Around 100 participants from Africa, Asia, Australia and Europe attended two keynote speeches, three workshops, 21 presentations, six poster presentations, 22 lightning talks and two special interest groups (SIGs). Results, developments and issues relating to STACK were presented and discussed in a total of 16 sessions. In addition, the supporting program with conference dinner provided ample opportunity for professional and private discussions with colleagues.

As in previous years, authors were able to submit a written contribution in addition to their presentations. After reviewing all submissions, we are pleased to publish 16 contributions on the following 136 pages. Unlike in previous years, however, this time the contributions will not be published individually, but together in a conference proceedings.

I would like to thank the organizing committee of the event as well as all active participants, authors, reviewers and supporters who together ensured an interesting and successful event. Let's continue the development of STACK and the associated improvement of teaching together!

Amberg, July 2024

Michael Weinmann

Table of content

1. **Automatic assessment of the geometric performance of basic operations with complex numbers: An example question using STACK and JSXGraph** 1

   Bernhard Gailer, Stephan Bach, Mike Altieri; OTH Amberg-Weiden

2. **Building Open STACK Question Banks** 17

   Emmaculate Odhiambo, Mary Sayuni; INNODEMS

3. **Data Procession Made Easy: A Python Tool for Extracting Information from Student Responses to STACK Questions** 26

   Jonas Lache; Hochschule Ruhr West and Ruhr-Universität Bochum

4. **Designing graphical physics problems with JSXGraph** 33

   Michael Kubocz, David Lauter, Stefan Roth; RWTH Aachen University, Department of Physics

5. **Enhancing Statistics Learning through Automated STACK Testing** 39

   Elena Safiulina, Natalja Maksimova, Oksana Labanova; TTK University of Applied Sciences

6. **GeoGebra in STACK. How it is done and what you need to consider for calculating feedback** 49

   Tim Lutz; University of Kaiserslautern-Landau

7. **Implementation and effectiveness of mathematical STACK problems in higher education for engineers: Lessons learned.** 54

   Oleg Boruch Ioffe, Maike Schelhorn, Jessica Schäfer, Reik V. Donner; h² – Hochschule Magdeburg-Stendal

8. **Investigating the feasibility of Automating the Advanced Higher Mathematics and Physics Scottish school exams** 59

   Konstantina Zerva, Ben Chooyin, Jack Fortune, Rachel Hollingsworth, Mitchell McLachlan, Heather Napthine; School of Mathematics, The University of Edinburgh

9. **New Developments Around the Evaluation Tool STACKrate** 75

   Daniel Meißner; Ruhr-Universität Bochum
   Jonas Lache; Hochschule Ruhr West and Ruhr-Universität Bochum

10. **On ongoing extensions of a comprehensive collection of mathematical STACK problems in German language** 81

Maike Schelhorn, Oleg Boruch Ioffe, Katharina Jaenicke, Reik V. Donner; Magdeburg-Stendal University of Applied Sciences

11. **On the Testing of Linear Algebra with STACK System** 87

Masumi KAMEDA; Freelance,
Mitsuru UDAGAWA; Nagoya University

12. **STACK Assessment supported flipped learning model of undergraduate Mathematics in African Universities: a theoretical exploration** 94

Motognon Wastalas d'Assise Dogbalou; Universita degli Studi di Trieste
Herine Otieno; EduHubAfric

13. **STACK for PDEs** 103

George Ionita, Laura Kobel-Keller; ETH Zurich
Florian Spicher; Universität Bern

14. **STACK in physics across the curriculum** 111

Abigail Pied, Alastair Palmer, Eric Yao; University of Glasgow

15. **STACKing Further with STACK-JS** 119

Sam Fearn; Durham University

16. **Using STACK Beyond Mathematics** 130

Maciej Tadeusz Matuszewski; Durham University

Focuses: Visual e-assessment, Implementation and usage of JSXGraph, Implementation of question features with STACK-JS, Feedback.

Article number: 01

Automatic assessment of the geometric performance of basic operations with complex numbers: An example question using STACK and JSXGraph

Bernhard Gailer*, Stephan Bach, Mike Altieri

OTH Amberg-Weiden

Abstract

This article introduces an assessment question on the geometric performance and understanding of basic arithmetic operations with complex numbers. The question is implemented using STACK and JSXGraph on a Moodle platform. Complex numbers are a topic with multiple starting points for geometrical interpretation. The presented question focuses specific arithmetic operations such as addition/subtraction, complex conjugation, or scaling. Based on the importance of geometric thinking for understanding mathematics and some theoretical background of visualization in science teaching, the authors briefly illustrate the conception of the question. This includes considerations on how to support the geometric focus. Subsequently, important aspects of the implementation, namely question variants and randomization, graphical input, grading and specific feedback, and general feedback are discussed in more detail. Special attention is given to particulars of the integration of JSXGraph, such as an effective transfer of randomized variables, or adaptions on the feedback-level using the new STACK-JS. These implicate suggestions for question authors with similar concerns. The article emphasizes the potential of the combination of STACK and JSXGraph for the assessment of mathematics in general and geometrical understanding and the connection of representations in particular.

* Corresponding author: b.gailer@oth-aw.de

1 Introduction

Geometry is an important representation of mathematics, also in topic areas that are not essentially geometrical. Geometric thinking can foster understanding (Atiyah, 2001), and the connection of different representations is an important concept in the didactics of mathematics (Dohle & Prediger, 2020) and closely connected to mathematical problem solving (Heinze et al., 2009). Therefore, both geometric thinking and the connection of representations should be an essential part of the *assessment* of mathematics as well. However, paper-based assessment entails different limitations here, namely expansive grading, hardly possible randomization, and no synchronous connection of representations. Dynamic geometry software (DGS) on the other hand, can provide rich learning experiences for the geometric representation of mathematics and especially better reflects its dynamic aspect (Soto-Johnson, 2013). Thus, the combination of a mathematical assessment system with DGS is a highly promising approach for learning and assessment. And here STACK[1] (as the assessment system) and JSXGraph[2] (as the DGS) come into play.

Support for JSXGraph has been added to STACK with version 4.2 in 2018 (STACK project, n.D.). Several articles have been published since dealing with the potential of this software combination for the assessment of mathematics and technical subjects (Mai & Meyer, 2019; Bach & Altieri, 2021; Kraska & Schulz, 2021; Hooper & Jones, 2023). Special attention is given to the specific use case of a given DGS construction manipulated by learners and evaluated by the assessment system. Wassermann and Rathmann (2024) refer to this as “visual e-assessment”.

In this article, we introduce a STACK question on the *geometric* performance of basic *arithmetic* operations with complex numbers, such as addition, complex conjugation, scaling, or multiplication. The question is part of the final quiz of a self-learning module on complex numbers that is designed for engineering programs, especially at universities for applied sciences. The self-learning module is implemented as a course in the learning management system Moodle.

We start with a short theoretical background on different aspects of visualization in mathematics, also addressing specific issues of complex numbers (section 2). After presenting the basic idea of the STACK question, including general didactical considerations (section 3), we delve into the implementation in section 4 . Concretely, we address the following aspects: 1. question variants and randomization, 2. graphical input, 3. grading and specific feedback, and 4. general feedback. Particular attention will be paid to the interplay between STACK and JSXGraph and adaptions on the feedback level using the new STACK-JS. The article closes with a brief discussion including lessons learned from the perspective of the developers.

2 Theoretical Background

Atiyah (2001) is considering geometry as – alongside algebra – one of “the two formal pillars of mathematics” (p. 657). He points out that, since geometry is essentially dealing with space and thus literally visible, it allows to use spatial intuition. This can support the comprehension of concepts and initiate deeper thinking.

But what do we mean by visualizing mathematical or in general scientific objects or concepts? There are three important aspects to the concept of visualization (Vavra et al., 2011): First,

[1] https://stack-assessment.org/
[2] https://jsxgraph.org/wp/

visualization objects simply refer to any kind of physical pictures of mathematical objects. These can be function graphs, geometrical sketches, or diagrams. In the context of complex numbers, an example of a visualization object is the representation of a complex number as a point or a position vector in the Gaussian Plane. Secondly, *introspective visualizations* are mental pictures. Thus, they denote how learners imagine mathematical objects. And last, *interpretive visualizations* name the process of making sense of visualization objects or introspective visualizations and connecting them to existing knowledge. While the first two aspects are nouns, interpretative visualization is a verb – a cognitive process (Soto-Johnson, 2014). Soto-Johnson (2014) points out that “in order for students to have a fully developed geometrical interpretation of a mathematical concept/notion/situation, they must have opportunities to experience both the noun and the verb cases of visualization” (p. 104). And here, DGS come into play, because in order to support a dynamic process such as interpretive visualization, it is helpful to have a dynamic system as well. DGS can much better represent students’ thinking than a static system or paper and pencil (Nelson, 2018).

Overall, we see that visualization is not only an important part of mathematics itself but also of teaching and learning the subject. Therefore, following the idea of constructive alignment (Biggs, 1996), it should part of mathematical assessment too, including interpretative visualization.

For the topic area of complex numbers Soto-Johnson (2014) lists five questions that are suitable for geometric interpretation: 1. definition of imaginary and complex numbers, 2. representations of complex numbers and in particular the imaginary unit, 3. addition and subtraction, 4. multiplication and division, and 5. any other arithmetic operation of complex numbers. Apparently, the arithmetic of complex numbers offers numerous starting points for geometrical interpretations. In the following section, we introduce a STACK question dealing with some of the questions above on a geometrical level.

3 STACK question on visualizing the arithmetic of complex numbers

The goal of the question introduced in this article, is to assess the geometric performance and understanding of basic arithmetic operations with complex numbers. This includes both introspective and interpretive visualization. The first is realized by giving symbolic expressions such as $z-\bar{z}$, that are to be drawn as geometric objects (points and position vectors) in the Gaussian Plane whereat the complex number z is already given as a position vector (see Figure 1). For in order to present the relevant visualization objects, learners first have to create mental images. The second is implemented by giving expressions that learners most likely don’t have introspective visualizations of yet. They might have mental images of a complex number z and its complex conjugate $\bar{z}$ but probably not of the difference $z-\bar{z}$. Thus, they have to make sense of the given expressions by connecting it to prior knowledge.

However, learners could get around interpretive visualization by just reading the cartesian coordinates of z from the applet, calculate $z-\bar{z}$ arithmetically, and again represent the result in the Gaussian plane. In order to prevent this, we have to forgo any numbers, such as grid, unit lengths, numbered ticks, or the symbolic expression of the given number z. If learners don’t have anything to calculate with they are forced to truly move to the geometrical level.

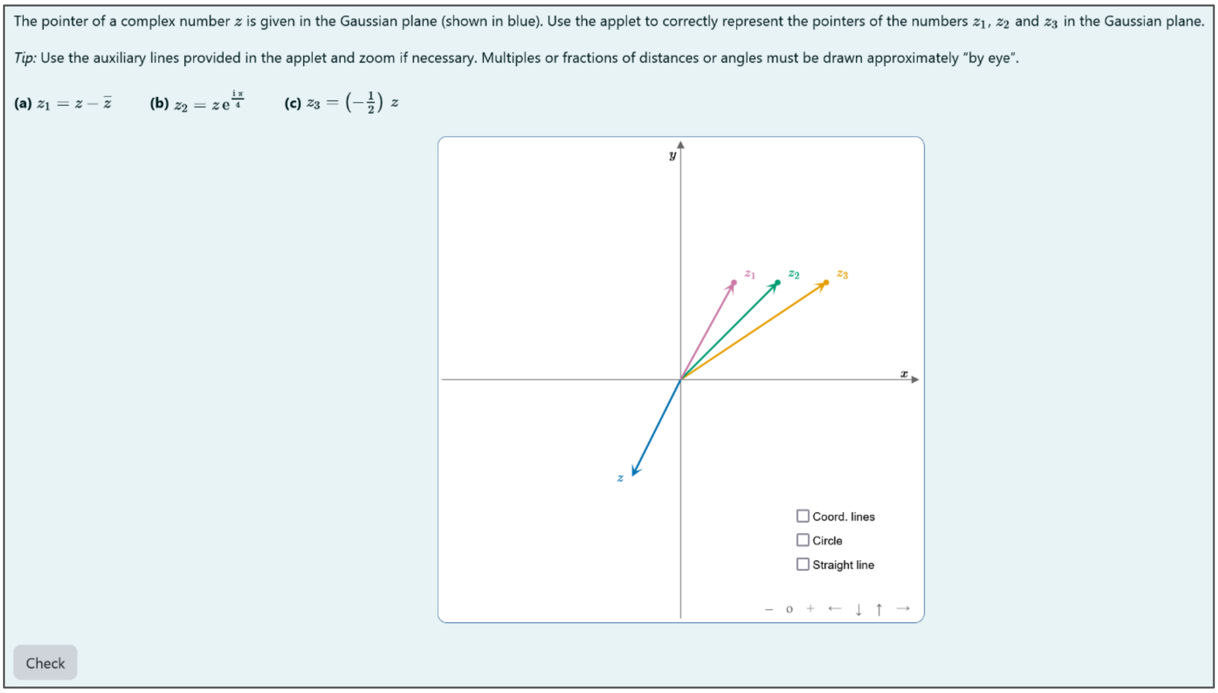


Figure 1: The presented STACK question on basic operations with complex numbers

This omission of numbers effects subtasks and randomization, because not every of the questions listed by Soto-Johnson (2014, see section 2) can be answered in this way. To geometrically represent roots and powers of a complex number, for example, would at least demand a unit length. We came up with three subtasks (a, b, and c), each focusing a different group of operations. The subtasks are randomized with all variants being structurally equal but not in view of the details of the operations occurring. Table 1 gives an overview of the different subtasks, variants, and corresponding operations.

Table 1: conception of the presented STACK question: Tasks, variants and corresponding operations

Task	Variants	Arithmetic operation(s)	Geometrical operation(s)
(a)	$z_1 = z - \bar{z}$ $z_1 = z + \bar{z}$ $z_1 = \bar{z} - z$	addition/subtraction; complex conjugation	translation; axial reflection
(b)	$z_2 = z \cdot i$ $z_2 = \frac{z}{i}$ $z_2 = z \cdot e^{\frac{i\pi}{4}}$	multiplication/division	rotation
(c)	$z_3 = -2 \cdot z$ $z_3 = -\frac{1}{2} \cdot z$ $z_3 = 2 \cdot \bar{z}$ $z_3 = \frac{1}{2} \cdot \bar{z}$	scaling; complex conjugation	stretching/compression, if applicable point reflection; axial reflection

4 Features and implementation

In this section, we present some of the main features of our question alongside with details of the implementation. For each feature, we start with a brief description of the motivation followed by possible challenges in the implementation. Finally, we present how we dealt with these challenges, in particular explaining the interplay between Maxima, JavaScript (in the form of STACK-JS), and STACK CASText.

Question variants and randomization

What do we want?

The main goal of using randomization and thus having different variants of the question is to prevent learners from cheating in exam situations and to provide multiple learning opportunities with structurally equal tasks. The latter was especially relevant in our setting because learners had two attempts on the final quiz. By having different variants, we could provide detailed feedback on the first test attempt without giving away the answer for the second try. In this way, learners can use feedback “to go the next step”.

Regarding our STACK question on visualizing arithmetic operations with complex numbers, we want to have different variants with the same difficulty but different operations to perform in each of the three tasks. Each variant offers structurally similar tasks with arithmetic operations of two complex numbers. These variants should be generated by randomizing certain parts with the help of Maxima.

What are possible challenges?

In order to enable the learning opportunities mentioned above, we cannot solely rely on the randomization of numbers. Instead, we have to randomize the whole arithmetic operation in each of the three subtasks. This confronts us with a challenge regarding the implementation of both the applet and the feedback. For there are different structures of the teacher’s answer in the subtasks, as well as different ways to get to this answer geometrically. For instance, subtask (b) can either be a multiplication or a division by the imaginary unit i depending on the question variant. The feedback thus has to be adapted to the concrete variant.

Another challenge regarding the randomization in our STACK question is the issue of visual feedback. While multiplication and division both result in a rotation on the geometric level, subtask (c) calls for a reflection only for some variants depending on whether the complex-conjugate is occurring or not. This must be considered when constructing a visual representation of the teacher’s answer.

How do we implement it?

Randomization starts with the definition of suitable question variables. First, we construct a randomized complex number z (Maxima variable `zcomp`) that is shown in the applet as a position vector. It is the basis for the operations given in the subtasks. The challenging part of randomization in our STACK question is the implementation of the different arithmetic operations for each subtask. We do this by first defining a random number, which is then used to pick an element from a list containing the different operations.

```
ta1_expr: [[z-zbar,zcomp -
conjugate(zcomp)],[z+zbar,zcomp+conjugate(zcomp)],[zbar-
z,conjugate(zcomp)-zcomp]][randnum1];
ta2_expr: [[z*%i,zcomp*%i], [z/%i,zcomp/%i],
[z*e^(i*%pi/4),zcomp*e^(i*%pi/4)]][randnum2];
ta3_expr: [[-2*z,-2*zcomp], [-(1/2)*z,-(1/2)*zcomp], [2*zbar,
2*conjugate(zcomp)], [(1/2)*zbar,(1/2)*conjugate(zcomp)]][randnum3];
```

Listing 1: Implementation of the randomized arithmetic operations for each task

Each element in the list is a list itself, containing one entry (the second one) for calculation on the Maxima side and another one for display purposes (the first one)[3]. A benefit of first constructing random numbers and then using them to pick elements from a list is that we can later identify which element and thus operation was picked. This would not be as easy if we directly used the function `rand()` with the list as its argument.

We include the randomly chosen symbolic expression in the question text by using CASText syntax, e.g. `\(z_1 = {@ta1_expr[1]@}\)`.

The teacher's answers for grading and feedback are defined in the question variables. Since the three complex numbers are to be represented as points and position vectors in the complex plane, the teacher's answer in each case is a list containing the cartesian coordinates. We later use these variables in the feedback variables to check the correctness of the student's answers.

```
ta1: [float(realpart(ta1_expr[2])), float(imagpart(ta1_expr[2]))];
ta2: [float(realpart(ta2_expr[2])), float(imagpart(ta2_expr[2]))];
ta3: [float(realpart(ta3_expr[2])), float(imagpart(ta3_expr[2]))];
```

Listing 2: Constructing the teacher's answers

Another challenging aspect regarding the randomization of STACK questions is the interplay between the Maxima variables and JSXGraph. As we want to visualize the randomized complex number z in the JSXGraph applet, we need to have access to our Maxima variables from within the code of the applet. For this, we include the Maxima variables in the JSXGraph code by using CASText syntax. In order to simplify the development process, we define a constant that holds all the Maxima variables that we need to access.

```
const MAXIMA = {
z_real: {#z_real#},
z_imag: {#z_real#},
ta1: {#ta1#},
ta2: {#ta2#},
ta3: {#ta3#},
…
};
```

Listing 3: Using variables from Maxima in the JSXGraph code

```
const zPoint = board.create("point", [MAXIMA.z_real, MAXIMA.z_imag],
{ name: "\\(z\\)", size: 0.1, showInfobox: false, fixed: true,
label: { offset: [-16, 0], color: JXG.palette.blue } });
```

Listing 4: Creating a point for the complex number z by using coordinates from Maxima

This allows to develop the question outside of the Moodle plain text editor as one can simply supply dummy values during development and later replace them with the corresponding CASText syntax for the actual variables. Also, one can quickly get an overview of the Maxima values used in the applet. In the JSXGraph code, one can use these values with normal JavaScript notation for accessing properties of an object (see Listing 4).

[3] We deviate from the display of the complex-conjugate in STACK in our self-learning module as we denote it with a bar: $\bar{z}$.

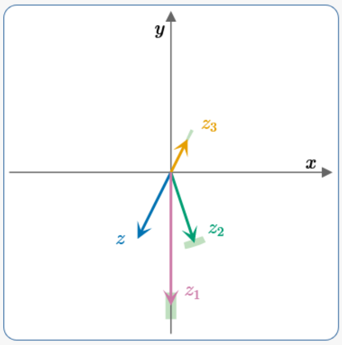
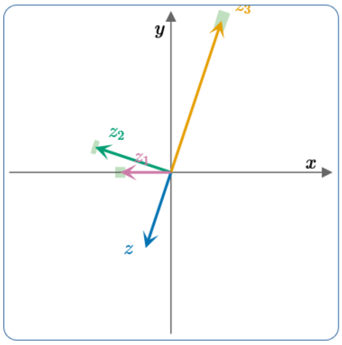
Variant	Question note
382745349	$z = -1 - 2\,\mathrm{i} \quad z_1 = z - \bar{z} \quad z_2 = z\,\mathrm{e}^{\frac{\mathrm{i}\,\pi}{4}} \quad z_3 = \left(-\frac{1}{2}\right)\,z$ 
525849564	$z = -1 - 3\,\mathrm{i} \quad z_1 = z + \bar{z} \quad z_2 = \frac{z}{\mathrm{i}} \quad z_3 = -2\,z$ 

Figure 2: Displaying the deployed question variants in the question

Randomization in STACK generates the need for a question note in order to distinguish the deployed variants. We opted to include a static (but randomized) copy of our main applet in the question note alongside the display of the given complex number z and the sought numbers z_1, z_2 and z_3. The applet is an integral part of the question and should be present in the question note. This allows both teachers and developers to get a quick overview of the display of different variants. We also display the tolerance areas for correct answers, which makes it easier to check if the tolerances that are implemented are appropriate.[4]

Graphical input

What do we want?

As elaborated in section 3 , in our question we want students to interpret and perform arithmetic operations solely on the geometrical level. Only using graphical inputs presents the learners with the need to move away from doing calculations on an algebraic level and instead tackle the problem geometrically. In addition, as explained in section 3 , we deliberately forgo some elements that can often be found in applets like this: grid, (numbered) ticks and measuring tools. In order to nonetheless draw the resulting complex numbers correctly, learners must use a mixture of construction and estimation. To support this process, we want to provide specific dynamic auxiliary lines: A set of coordinate lines, a line through origin and an auxiliary circle that helps preserving the absolute value when performing rotations or reflections.

[4] The implementation of the tolerance areas and their use for feedback purposes will be explained in detail in the subsection on grading and specific feedback.

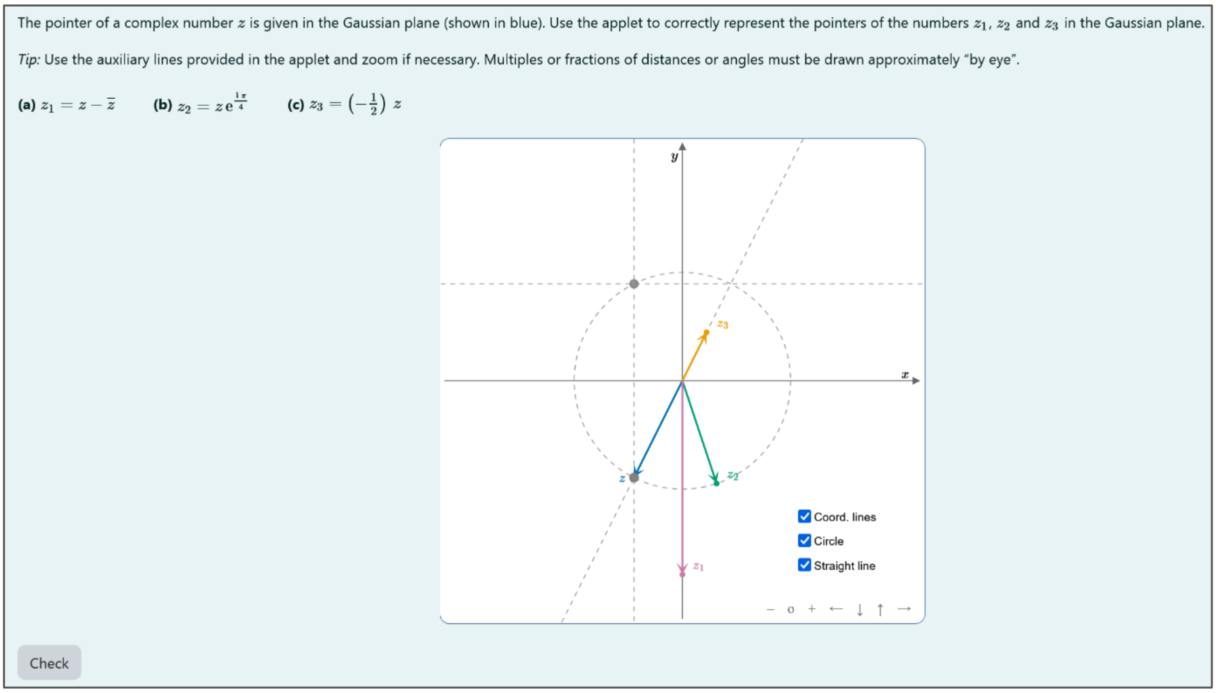


Figure 3: Graphical Input with the correctly constructed answers using auxiliary lines

What are possible challenges?

Because we want learners to construct their solution graphically inside an applet, we must provide a way for them to move the corresponding points in the complex plane. Moreover, the cartesian coordinates of these points must be stored and kept in sync with the positions in the applet. The state of the applet needs to be remembered after the question has been submitted, so we must make use of data bindings between STACK and JSXGraph.

How do we implement it?

We chose to initially display the position vectors of the complex numbers z_1, z_2 and z_3 at arbitrary positions in the JSXGraph applet. Learners must then move the tip of the vectors to the correct positions. We opted for this approach as it would be more expansive to provide an option for learners to "draw" a number on the complex plane or to use buttons for adding a number. Also, in this particular question, it is obvious that exactly three numbers are to be presented.

In order for the auxiliary lines to be actually helpful, they must be easy to place at certain positions. For instance, the helper line that can be used to perform an elongation in subtask (c) (e.g. $z_3 = -2 \cdot z$) snaps to the position of z when dragged there. This is done by defining z as an attractor of the dragging point of the auxiliary line. A similar feature is implemented for the auxiliary circle.

```
helperLine.point2.setAttribute({ attractors: [zPoint, helperCircle],
attractorDistance: .1, snatchDistance: .3 });
helperCircle.point2.setAttribute({ attractors: [zPoint, helperLine],
attractorDistance: .1, snatchDistance: .3 });
```

Listing 5: Defining attractors for the points of the auxiliary objects

The graphical input is implemented using hidden input fields in the question text that are linked to the JSXGraph applet. This is done by using binding functions from the stack_js namespace. In order to get the coordinates of the three points that learners have to position, we use the function `stack_jxg.bind_point`. Also, the hidden inputs receive the points' initial coordinates by means of the function `stack_jxg.starts_moved`. This ensures that learners get feedback for all three tasks regardless of whether they moved the corresponding point or not.

```
stack_jxg.bind_point(ans1Ref, z1Point);
…
stack_jxg.starts_moved(z1Point);
…
let inputBinderZoom = function (inputRef) {
let serializer = () => JSON.stringify(board.getBoundingBox());
let deserializer = function (data) {
board.setBoundingBox(JSON.parse(data));
}
stack_jxg.custom_bind(inputRef, serializer, deserializer, [xAxis,
yAxis])
}
inputBinderZoom(ans4Ref);
```

Listing 6: Binding functions in the JSXGraph code

Another binding is implemented for remembering the bounding box of the applet. This is useful because in some scenarios, learners must zoom out a little bit in order to construct their solution. The zoom level is thus preserved when the question is submitted or learners re-enter the quiz. This binding is done by the newly added function `stack_jxg.custom_bind` that allows for a more advanced binding logic. It is ideal for scenarios like this, where you want to store state or, more generally, where you move beyond just tracking point coordinates.

Grading and specific feedback

What do we want?

We want to evaluate and grade learners' answers (i.e. position vectors in the JSXGraph applet) and also provide feedback indicating if an answer was correct, partially correct, or incorrect. The specific feedback should be closely linked to the graphical input, as feedback and graphics that belong together should be placed close to each other. This is why, in addition to standard symbolic feedback next to the algebraic display of the three expressions, we opted for feedback inside the applet, i.e. close to the actual answer. This kind of spatial contiguity is beneficial in multimedia learning (Mayer, 2009).

Since learners cannot determine (by construction or calculation) exact positions of the complex numbers they also must rely on estimations. In this context, textual feedback or standard symbolic feedback in the form of check- or crossmarks has the downside that it can't give learners an idea of how close their solution is to the correct answer. For instance, a learner constructed the number $z_3 = \frac{1}{2}\,\bar{z}$ and got the angle correct but not the length of the position vector. Now, when using textual feedback, one can only say something like "the length of the vector is almost correct" but it is non-transparent for the learner how much his solution deviates from the teacher's answer. In order to mitigate this issue, we opted to provide feedback in form of a coloured area around the correct solution. Learners can now see how close their answer is to the teacher's answer which, especially in the case of partially correct answers, motivates to engage with their own solution.

As for grading, we want to give credits as closely as a human grader. I.e. answers that are quite close to the teacher's answer should receive partial credit and answers that are very close to the teacher's answer full credit. Also, if one property of the answer can be determined more precisely than another it should be evaluated more strictly.

What are possible challenges?

Since we rely on graphical input and deliberately renounce grid lines and measurement tools, we cannot expect the student's answer to be fully accurate. Therefore, tolerances for the cartesian and/or the polar coordinates – depending on the demands of the subtask – are needed. These tolerances are not evident by default and randomization makes them especially tricky to determine: In some variants the absolute value is greater than in others, in the first subtask the teacher's answer is always positioned on one of the axes, if learners us the auxiliary circle, the length of the position vector can be drawn exactly, …

Furthermore, we let learners construct their solutions geometrically but cannot assess in the same way. We are limited to numbers i.e. cartesian or polar coordinates, which creates a challenge on the technical level: Learners should answer geometrically but we evaluate numerically.

How do we implement it?

We define the tolerances needed in the feedback variables of the PRTs. For instance, in the second task learners must perform a multiplication or division with a complex number with the magnitude 1 which results in a rotation of z. In an algebraic context this operation would typically be performed using the polar form of complex numbers. We opt to evaluate the answers accordingly and thus check the polar coordinates (absolute value and argument) of the position vector of z_2. For answers to be graded as correct, there is a small absolute tolerance of the absolute value, because the provided helper circle allows a pretty exact construction. As for the argument of z_2 we opted for a medium absolute tolerance depending on the rotation angle because learners need to estimate this part. This results in the tolerance area being a sector on a circle ring. For marking answers partially correct, we opted for a relative tolerance of minimum 0.3 and up to 15% of the distance to the correct point, i.e. the tolerance area is a circle in this case.

```
tolb_carg: if turnb > 45 then 5 else 8;
tolb_carg: ev(float((tolb_carg*%pi)/180),simp);
tolb_cabs: 0.1;
partcorrb_dist: max(0.3, 0.15*Distance(ta2, [0,0]));
```

Listing 7: Tolerances for grading answers for z_2 in the second task of the presented STACK question

We then define our feedback variables using these tolerances. First, we construct the Cartesian form of the answer and then define predicates for correct and partially correct answers using Maxima's functions `cabs` and `carg`. These predicates are used in the PRT nodes with the CasEqual[5] answer test.

[5] Since the feedback variables for correctness (e.g. `z2_corr`) are evaluated to either `true` or `false`, CasEqual is the simplest answer test one can use. Of course, AlgEquiv would be possible as well.

```
ans2_z2_complex: ans2_z2[1]+%i*ans2_z2[2];
/* Correct */
z2_corr: is(abs(cabs(ans2_z2_complex)-cabs(ta2_expr[2])) <=
tolb_cabs) and is(abs(float(carg(ans2_z2_complex)-
carg(ta2_expr[2])))) <= tolb_carg);
/* Partially correct */
z2_part_corr: is(Distance(ans2_z2, ta2)<= partcorrb_dist);
```

Listing 8: Feedback variables for grading answers for z_2

In Listing 8, one can see assessment process described earlier: Initially, we present an algebraic representation of an expression. Learners visualize this expression geometrically and we use the point coordinates to construct the algebraic representation of their answer in order to evaluate. This applies to all subtasks.

The Maxima-code in Listing 8 is not only used for assessment and grading but also for displaying graphical feedback inside the applet. This can be done by placing HTML span elements with unique IDs inside the PRT nodes. In the JSXGraph code, we use the STACK-JS function `get_content` that lets us look outside of the iframe in which the JSXGraph code is executed. This is an asynchronous operation so we have to await the result and if a span with the corresponding ID is returned, we can be sure that the feedback we are looking for is present in the STACK question. With this approach, we can draw the tolerance area for correct and partially correct answers inside the applet when the corresponding PRT feedback is present. For subtask (b), for example, this results in a green sector for correct answers (tolerances for the absolute value and the argument of the polar form) and an orange circle (tolerance for the distance to the correct point) for partially correct answers. Listing 9 alongside Figure 4 shows the JSXGraph Code together with the resulting graphical feedback.

```
// z2 correct feedback
stack_js.get_content("A2-z2_correct").then(content => {
if (content !== null) {
drawCorrectSector(MAXIMA.ta2, MAXIMA.tolb_carg, MAXIMA.tolb_cabs);
} });
// z2 partially correct feedback
stack_js.get_content("A2-z2_partially_correct").then(content => {
if (content !== null) {
board.create("circle", [[MAXIMA.ta2[0], MAXIMA.ta2[1]],
MAXIMA.partcorrb_dist], { name: "\\(z_2\\)",
...correctPointAreaStyle, fillColor: "orange" });
} });
```

Listing 9: JSXGraph code for checking for feedback in the STACK question and showing the graphical feedback

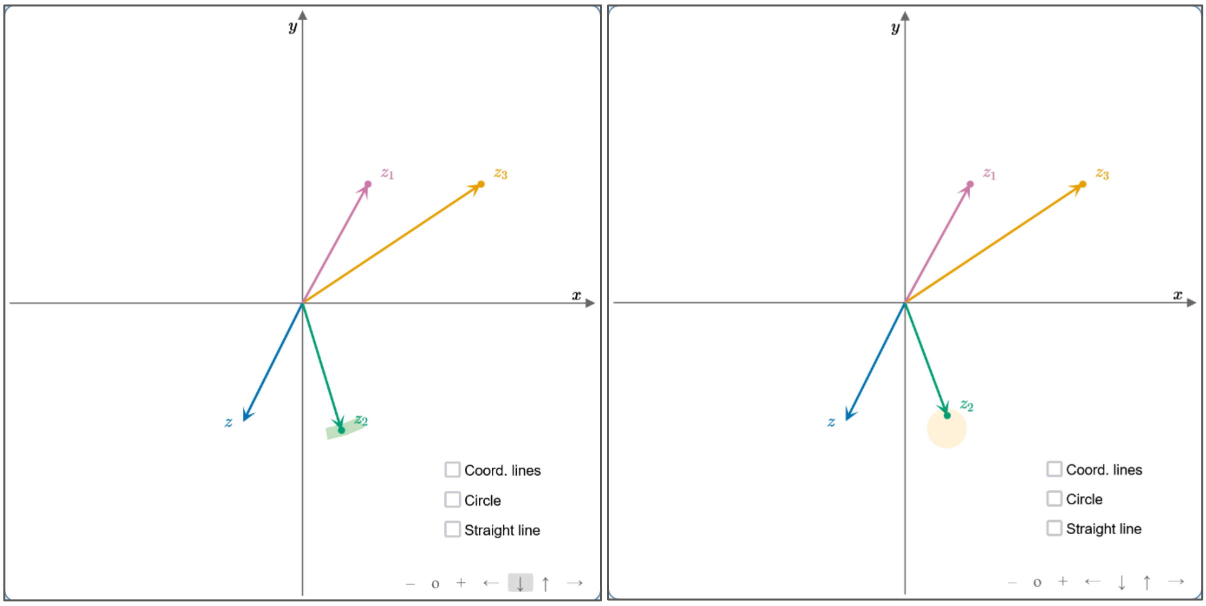


Figure 4: Graphical feedback for a correct (left hand) and a partially correct (right hand) answer for z_2

In addition to the specific feedback for each subtask, we implemented a formative PRT that gives a motivational feedback to learners who got all three subtasks correct. In other cases, it gives a short information if at least some or none of the answers are correct. Instead of using multiple PRT nodes, we opted for a single node that just checks if all answers are correct or not. In the false section of this node, we then make use of conditional blocks to check for the aforementioned details. This approach makes the PRT structure much simpler but also comes with limitations especially when testing the PRT. However, for formative PRTs which do not contribute to grading, this is a veritable way of keeping the PRT small and concise.

```
[[ if test='is(some_corr)' ]]
<p>You have not yet drawn all the numbers correctly.<p>
[[ elif test='not(some_corr) and is(some_partcorr)' ]]
<p>You have not drawn any of the three numbers correctly. However,
the numbers you have drawn are {@if is(z1_part_corr and z2_part_corr
and z3_part_corr) then "" else "partly"@}  already close to the
correct solutions.<p>
[[ else ]]
<p>You have not drawn any of the three numbers correctly.<p>
[[/ if ]]
```

Listing 10: Formative feedback using conditional blocks and inline Maxima if-else statement to provide information on the answers

General feedback

What do we want?

With the help of the general feedback, we want learners to reflect on their performance and retrace the correct way of solving each task. Therefore, the general feedback is a fitting section to include a worked solution. As learners must solve this question geometrically using an applet, we opted to not only include a worked solution in the form of text but also in the form of another applet. This applet shows the constructions for the correct answers. Our goal behind presenting learners the worked solution using two different display modes is to foster deeper understanding by connecting different representations (see sections 1 and 2).

What are possible challenges?

Since our randomization goes beyond values for numbers, we have to be very careful about providing a comprehensible feedback no matter what variant is shown to the learners. This is a particular challenge for the implementation of the applet. Because there are some variants of a task, for example, that include the complex conjugate of z and thus require a reflection together with scaling while this is not the case for others. The construction shown in the applet must only display the necessary steps as showing helper lines that are not required would lead to confusion on the learner's side. For instance, we do not want to show the helper circle if the correct answer for the task is only constructed by translation and reflection.

How do we implement it?

The implementation of the textual part of the general feedback is pretty straight-forward. To account for the randomization, we again make use of inline Maxima if-else statements inside CASText delimiters. This allows for writing a single paragraph inside the editor that displays different information based on the question variant. As our sentence will always have the same structure, we do not need to include conditional blocks here, but they would be very helpful in scenarios where one wants to display truly different content depending on randomization.

```
<!-- Feedback z2 -->
<p style="margin-top: 1em;">
<b>(b)</b>  The pointer of \(z_2 = {@ta2_expr[1]@}\) is
obtained in the Gaussian plane by rotating the pointer of \(z\) by
\({@if is(randnum2<=2) then 90 else 45 @}^\circ\) in the
mathematically {@if oddp(randnum2) then "positive" else "negative"@}
direction of rotation, i.e. {@if oddp(randnum2) then
"counterclockwise" else "clockwise" @}{@if evenp(randnum2) then "
(or by \\(270^\\circ\\) counterclockwise)" else ""@}.
</p>
```

Listing 11: General feedback for the second task using CASText delimiter with Maxima if-else statements to account for random variants

As for the JSXGraph applet in the general feedback, things get a bit more complicated. We now have to display the teacher's answer along with the relevant helper elements in the correct position to display the construction of the teacher's answer. As pointed out earlier, the number of helper elements to display as well as their position varies from variant to variant, so we cannot hard code the solution in the JSXGraph code. Instead, we first determine which variant numbers require which helper elements and based on that classification we transfer the display information from Maxima to JSXGraph alongside the teacher's answer. Therefore, we use Maxima code in the question variables to generate lists containing coordinates for the positioning of the auxiliary lines as well as in some cases a corresponding label.

```
/* Helper vectors */
MAXIMA.helpers_a.forEach(helper => {
if (helper.length) {
helpersaVecs.push(board.create("segment", [[0, 0], [helper[0],
helper[1]]], { ...pointerStyle, strokeColor: 'darkgray', name:
helper[2], withLabel: true, visible: true, label: { color:
"darkgray", autoposition: true } }));
}
});
```

Listing 12: JSXGraph code using variables from Maxima to conditionally draw helper elements for the construction of subtask (a)

These variables are then used in the JSXGraph code to create the construction for each task with respect to the variant. For example, we use `MAXIMA.helpers_a` to create a vector for the complex conjugate of z or the negative of the complex conjugate. The result is an applet that provides the construction for each of the subtasks via radio buttons.

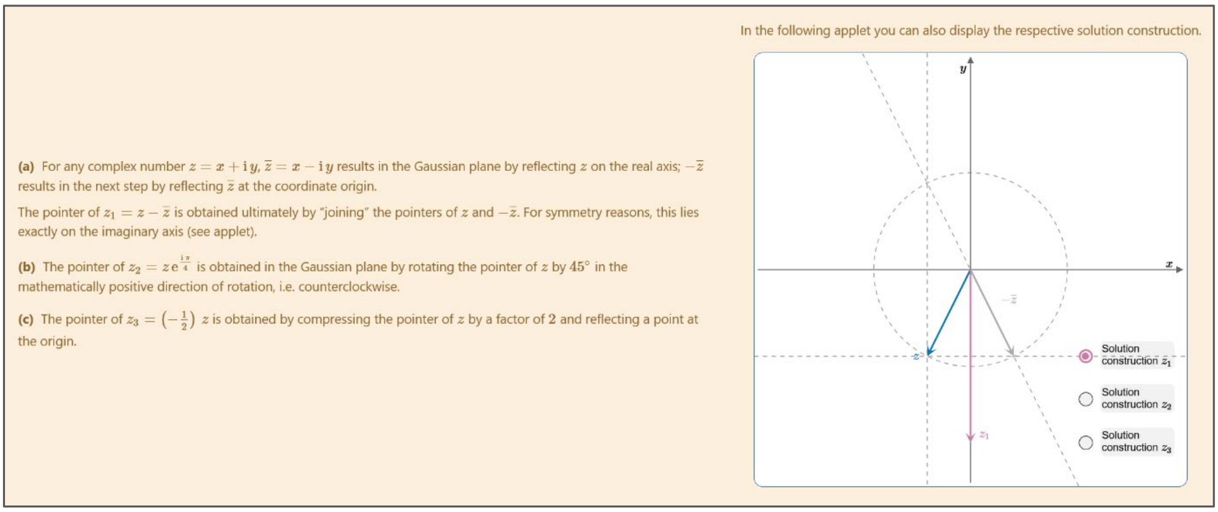


Figure 5 General feedback applet showing the construction of the teacher's answer for each task

5 Discussion

Overall, one can say that combining STACK with JSXGraph gives many options to incorporate visualization of mathematics into assessment including visual e-assessment for summative and formative reasons. The arithmetic of complex numbers is particularly well-suited for this concern. One strength of STACK and JSXGraph combined is the possibility to use a variety of input types which can lead to more engagement and a deeper understanding of the topic. Another advantage is the provision of various features to support feedback and interactivity, such as, for instance, reacting to feedback from the potential response trees in the applet or storing graphical input information for algebraic assessment with Maxima.

In view of the technical implementation, there have been quite a lot of changes in the last years regarding the way STACK handles client-side JavaScript. The new STACK-JS provides many useful functions for the integration of additional features into STACK and JSXGraph. For instance, STACK-JS allows to check if feedback is present in the question and then change the display of your JSXGraph visualization as you can not only use STACK-JS within the new JavaScript blocks but also seamlessly inside your JSXGraph blocks.

Randomization of STACK questions with JSXGraph that goes beyond picking numbers for variables and also concerns the structure of expressions can be challenging and time-consuming. This mainly regards the implementation of feedback and worked solutions, because in such cases their structure often differs as well. However, with feedback variables, conditional blocks, and inline Maxima code STACK offers various features to also handle extensive randomization.

After developing the presented question, we see some issues for future considerations, especially regarding didactics and trial with students. The first aspect concerns the use of auxiliary lines and the weighing between construction and estimation. Currently, learners have to use a mixture of both. For instance, providing more auxiliary lines would change the workflow towards a more construction-based and thus more algorithmic process, which might reduce the use of intuition. Secondly, up to now there has only been a trial of an earlier version of the question. Trials and regular use of the current version are in planning. This will allow to evaluate in which subtasks or variants learners typically face problems and – based on this – to add (or remove) auxiliary lines, adapt tolerances, and improve feedback.

Acknowledgments

The publication is part of the project “Innovationsnetzwerk für digitale adaptive Lehre” (acronym IdeaL), which is funded by the Stiftung Innovation der Hochschullehre from 2021–2024 (funding code FBM2020-EA-2870-07550, project leader Mike Altieri, OTH Amberg-Weiden).

Bibliography

Atiyah, M. (2001). Mathematics in the 20th Century. *The American Mathematical Monthly*, *108*(7), 654-666.

Bach, S., & Altieri, M. (2021). Drawing Graphs of Differentiable Functions with STACK and JSXGraph using Hermite Splines. *International Meeting of the STACK Community 2021.* https://doi.org/10.5281/zenodo.5035977

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher education, 32(3)*, 347-364.

Dohle, A., & Prediger, S. (2020). Algebraische Terme durch Darstellungsvernetzung und Scaffolding verstehen – Einblicke in einen sprachbildenden Mathematikunterricht. *Sprachförderung und Sprachtherapie in Schule und Praxis, 9*(1), 16–24.

Heinze, A., Star, J. R., & Verschaffel, L. (2009). Flexible and adaptive use of strategies and representations in mathematics education. *ZDM*, *41*, 535-540.

Hooper, C., & Jones, I. (2023). *Conceptual statistical assessment using JSXGraph.* International Association of Online Engineering. https://doi.org/10.3991/ijet.v18i01.36529

Kraska, M., & Schulz, D. (2021, Juni 9). Automatic assessment of free body diagrams using STACK. *Contributions to the International Meeting of the STACK Community 2021.* https://doi.org/10.5281/zenodo.5083692

Mai, T., & Meyer, A. (2019). Sketching functions as a digital task with automated feedback. *Contributions to the 1st International STACK Conference 2018.* https://doi.org/10.5281/zenodo.2582427

Nelson, M. (2018). Effects of Dynamic Geometry Software on Secondary Students' Understanding of Geometry Concepts. *Culminating Projects in Teacher Development* 36. https://repository.stcloudstate.edu/ed_etds/36

Richard E. Mayer (2009). *Multimedia Learning: Vol. 2nd ed*. Cambridge University Press.

Soto-Johnson, H. (2014). *Visualizing the arithmetic of complex numbers. International Journal of Technology in Mathematics Education, 21*(3), 103-114.

STACK project (o.D.). *STACK Development History*. stack-assessment.org. https://docs.stack-assessment.org/en/Developer/Development_history

Vavra, K. L., Janjic-Watrich, V., Loerke, K., Phillips, L. M., Norris, S. P., & Macnab, J. (2011). Visualization in science education. *Alberta Science Education Journal*, *41*(1), 22-30.

Wassermann, A., & Rathmann, W. (2024, March 11–13). *Visual e-Assessment with JSXGraph in Calculus* [conference presentation]. International Meeting of the STACK Community 2024, Amberg, Germany.

Information on the authors

Bernhard Gailer

Bernhard Gailer is a research assistant for digital tasks at OTH Amberg-Weiden. He first came into contact with STACK in 2019 during his business studies as a student assistant. After successfully completing his Master's degree at the OTH Amberg-Weiden, he has been working as a research assistant in the IdeaL project at the university since the beginning of 2022. There he is responsible for the development of STACK questions for interactive maths learning modules in Moodle. His main focus is on the integration of JSXGraph and STACK and the adoption of STACK-JS.

Stephan Bach

Stephan Bach studied math and physics for teaching and worked as a teacher for these subjects for several years. Since 2014 he is a research assistant at OTH Amberg-Weiden. There he has been involved with STACK in various projects and across subjects almost from the beginning on. He is especially interested in using STACK to improve conceptual understanding and to provide activating and specific feedback.

Mike Altieri

Mike Altieri studied mathematics and received his doctorate in mathematics didactics. From 2016 to 2019, he was Professor of Applied Mathematics and Didactics of Mathematics and Natural Sciences at the Ruhr West University of Applied Sciences. Since 2019, he has been Professor of Media Didactics at OTH Amberg-Weiden, where he heads the Competence Center for Digital Teaching and is Vice President for Teaching, Didactics, Digitization.

Focuses: STACK in teaching or exams; Other topics related to STACK.

Article number: 02

Building Open STACK Question Banks

Emmaculate Odhiambo*, Mary Sayuni

INNODEMS

Abstract:

The adoption of STACK in Kenyan institutions has been driven by the need for efficient assessment solutions, limited resources and increasing student enrolment. This has been implemented by collaboration among four mathematics and mathematics education graduates based at Maseno, Kenya trained and mentored by IDEMS International colleagues and lecturers from African Universities. The team developed, and is constantly improving, open question banks for undergraduate courses used mostly but not exclusively in universities across Africa, which are freely accessible to anyone. The collaborative efforts among the broader team have significantly contributed to the successful development and implementation of said Open Question Banks which not only support individual learning experiences but also facilitate easy access to diverse range of questions fostering a dynamic and resourceful education environment.

* Corresponding author: emmaculate.atieno@innodems.org

1. Introduction

In today's rapidly evolving educational landscape, the demand for effective assessment solutions is more pressing than ever, especially in Africa where educational resources are often constrained, yet student enrollment continues to rise. Traditional assessment methods, reliant on static question formats and manual grading processes, struggle to meet the diverse needs and learning styles of a growing student population. Recognizing these limitations, educators and institutions have increasingly turned to technology-driven solutions, such as the STACK (System for Teaching and Assessment using a Computer Algebra Kernel) system.

The growing interest in building open STACK question banks has been primarily driven by the system's unique capability to address challenges associated with limited resources, large class sizes, and the demand for effective assessment solutions that enhance student learning. These Open STACK question banks are freely accessible, hosted on IDEMS servers, providing African universities with unprecedented opportunities to overcome these challenges and elevate the quality of mathematics education for undergraduate courses across the continent.

This collaborative effort is supported by four mathematics and mathematics education graduates from Maseno University, Kenya, who have been trained and mentored by IDEMS International colleagues and lecturers from African universities. Together, they have played a pivotal role in developing and continuously improving these open STACK question banks, ensuring their relevance and accessibility to educators and students alike.

Throughout this article, we will delve into the development process of these open STACK question banks, their impact on student learning outcomes, challenges encountered during implementation, and strategic insights for sustaining and expanding this transformative initiative in African universities.

2. Development of Open STACK Question Banks

The development of open STACK question banks for African universities has been a collaborative endeavour aimed at enhancing the quality of mathematics education amidst challenges such as limited resources and increasing student enrolment. Spearheaded by mathematics and mathematics education graduates from Maseno University, Kenya, in collaboration with IDEMS International and lecturers from African universities, this initiative has been marked by innovation and dedication to improving assessment practices.

Training and Mentorship

At the core of the initiative's success lies the intensive training and mentorship provided to the mathematics graduates involved since July 2022. This training, conducted by IDEMS International in collaboration with African university lecturers, aimed to impart specialized skills in question authoring and technical implementation of the STACK system. To ensure a comprehensive understanding and proficiency in STACK, the team actively engaged in two weekly remote meetings and practical sessions led by the IDEMS team from the UK and Austria, as well as two week-long, in-person sprints. This training equipped the graduates with the expertise needed to develop adaptive assessment questions tailored to the requirements of undergraduate mathematics courses across Africa, consequently enriching the Open Stack Question Bank.

Areas of Proficiency:

- Technical Mastery: Proficiency in utilizing the STACK system, including its capabilities for generating dynamic and personalized assessments.
- Instructional Design: Ability to structure questions effectively to assess various levels of understanding and cognitive skills among undergraduate students.
- Pedagogical Expertise: Competence in integrating technology-enhanced learning experiences into mathematics education, ensuring alignment with curriculum objectives and educational standards.

Process of Question Creation

The development of Open STACK question banks began with identifying key learning objectives and curricular requirements across participating universities. Collaborative workshops and brainstorming sessions facilitated by IDEMS International enabled the team to design a comprehensive framework for question development.

During this phase, the team focused on:

- Content Development: Creating a wide range of mathematical questions that covered core concepts and advanced topics relevant to undergraduate mathematics courses.
- Quality Assurance: Conducting rigorous reviews and testing to ensure the accuracy, clarity, and educational value of each question.
- Accessibility Considerations: Designing the question banks to be user-friendly and accessible across different learning environments and technological capabilities present in African universities.

The team created questions generated by algorithms on the STACK platform, capable of offering varied follow-up questions or feedback based on students' prior responses. This adaptive method guarantees personalized feedback for each student, nurturing a profound grasp of mathematical concepts and improving overall learning outcomes.

Throughout the question creation process, each query undergoes a meticulous review to uphold its quality. This begins with a peer review, where feedback is solicited and any necessary improvements are noted. Upon receiving feedback, the author revises the question accordingly. Subsequently, the question undergoes a second review, following the same process as before to ensure its accuracy and effectiveness. This iterative development approach allows for continuous incorporation of feedback and enhancements, refining the question banks based on real-world usage and evolving educational practices. Feedback from educators and students who piloted the question banks provides valuable insights, enabling the team to address specific learning needs and enhance the overall effectiveness of the resources.

Quality Assurance and Continuous Improvement

Quality assurance measures were integrated throughout the development process to maintain the rigor and relevance of the question banks. Peer reviews, pilot testing in diverse educational settings, and feedback mechanisms from students and educators were instrumental in refining question formats and ensuring alignment with educational standards across African universities.

Continuous improvement remains a cornerstone of the initiative, with regular updates and revisions to the question banks based on ongoing feedback and advancements in educational technology. This iterative process ensures that the Open STACK Question Banks evolve in tandem with educational needs and technological advancements, thereby maximizing their impact on student learning and engagement.

3. Courses with Open STACK Question Banks

A pivotal decision in the project was to make the STACK question banks freely accessible. This choice stemmed from a commitment to advancing educational equity across Africa. By providing these resources at no cost, the initiative aimed to break down barriers to quality education and empower institutions with limited resources to elevate their teaching and assessment standards.

The Open Question Banks (OQBs) covering a wide array of mathematical disciplines play a crucial role in supporting education across various levels and topics within mathematics. Here's an analysis of how these question banks relate to specific courses:

1. Linear Algebra:
    - Topics Covered: Vector spaces, matrix theory, eigenvalues, eigenvectors etc.
    - Educational Value: Facilitates in-depth understanding and practice of foundational concepts in linear algebra.
2. Statistics and Probability:
    - Topics Covered include: Descriptive statistics, inferential techniques, data and analysis methods among others.
    - Educational Value: Provides a range of questions from basic to advanced levels, helping students grasp statistical concepts and develop skills in analysing and interpreting data.
3. Basic Mathematics:
    - Topics Covered: Arithmetic, algebra, geometry, number theory etc.
    - Educational Value: Offers fundamental practice in core mathematical skills essential for all disciplines.
4. Complex Numbers:
    - Topics Covered: Complex functions, contour integration, residue calculus among others.
    - Educational Value: Supports advanced mathematical studies by covering complex analysis, a critical area in mathematics. It enhances students' ability to understand and apply complex functions in applied contexts.
5. Advanced Mathematics:
    - Topics Covered: Broad range of advanced topics etc.
    - Educational Value: Includes higher-level concepts beyond basic coursework, catering to students pursuing deeper theoretical knowledge or specialized studies in mathematics.

6. Applied Mathematics:
    - Topics Covered: Differential equations, numerical methods, optimization etc.
    - Educational Value: Focuses on practical applications of mathematics .It prepares students to solve real-world problems using mathematical models and computational methods.
7. Calculus:
    - Topics Covered: Calculus concepts including differentiation, integration, and applications etc.
    - Educational Value: Essential for understanding rates of change and accumulation in various scientific and engineering disciplines. It provides foundational skills for advanced studies in mathematics and related fields.

4. Implementation of the Open Stack Question Banks

Upon completion, the Open Stack Question Banks were implemented within Maseno University, Masinde Muliro University of Science and Technology, Bahir Dar University Ethiopia, University of Namibia, Rongo University and The Technical University of Kenya among others. The positive reception and impact observed within these initial implementations paved the way for broader adoption across African universities. Educators reported enhanced engagement and learning outcomes among students, facilitated by the interactive and personalized nature of STACK-generated assessments.

5. Impact on Student Learning Outcomes

The implementation of Open STACK Question Banks in African universities has yielded significant improvements in student learning outcomes and educational experiences. This section explores the transformative impact of these adaptive assessment tools on undergraduate mathematics education across the continent.

Enhanced Student Engagement

One of the primary benefits of Open STACK Question Banks is their ability to enhance student engagement. By presenting dynamically generated questions that adapt to individual learning progress, the STACK platform promotes active participation and deeper interaction with course materials. Students are actively involved in problem-solving processes, which fosters a deeper understanding of mathematical concepts and improves retention rates.

Personalized Feedback Mechanisms

Central to the effectiveness of STACK is its personalized feedback mechanism. Through algorithmically generated questions, students receive immediate feedback based on their responses, identifying areas of strength and areas needing improvement. This tailored feedback not only enhances learning efficiency but also empowers students to take ownership of their academic progress and seek targeted support where needed.

Improvement in Learning Outcomes

Empirical evidence from universities implementing STACK has shown tangible improvements in learning outcomes. Students exposed to questions from the Open STACK Question Banks demonstrate higher levels of proficiency in mathematical problem-solving, critical thinking, and application of theoretical concepts to real-world scenarios. This enhanced competency prepares students for future academic pursuits and professional careers in STEM fields.

Accessibility and Equity

The free accessibility of Open STACK Question Banks hosted on IDEMS servers ensures equitable access to high-quality educational resources across African universities. Regardless of institutional resources or geographic location, students and educators can benefit from standardized, adaptive assessment tools that promote fairness and inclusivity in education.

Institutional Adaptation and Support

The integration of STACK into institutional frameworks has facilitated adaptive teaching practices and curriculum development. Educators have leveraged the platform's analytics to tailor instructional strategies, identify learning gaps, and personalize learning experiences for diverse student cohorts. This institutional adaptation enhances overall teaching effectiveness and student satisfaction with course offerings.

Case Studies and Success Stories

Several case studies from Maseno University, Masinde Muliro University of Science and Technology, Bahir Dar University Ethiopia and other African institutions highlight the transformative impact of Open STACK Question Banks on educational outcomes. These success stories underscore the scalability and efficacy of STACK in addressing the educational challenges faced by diverse student populations in Africa.

Here are the success stories:

> *"In 2018, I returned to Maseno after six years of PhD studies in the USA. One of my new responsibilities was teaching undergraduate maths classes and was immediately overwhelmed by the task. My first class was an introduction to Linear algebra that had 1000 students and I did not have any support in terms of teaching assistants. Clearly, giving homework, marking it, and giving immediate personalized feedback was going to be impossible, unless we innovated. The first time teaching the course, I used the traditional paper-based continuous assessment and I had a challenge marking. The next time I taught the course, I used STACK questions and was able to give weekly homework assignments and my students received immediate feedback which promoted learning on their part. Most students were very happy with the intervention and some expressed their hope of having more courses taught and their formative assessment done using STACK. Some students thought that it was too much work since they were used to only a few questions a semester in the form of continuous assessment. On my part as the course lecturer, I was greatly relieved since I could monitor student progress on a weekly basis and tailor my lectures accordingly. Formative assessment was easier and the exam performance for students that used STACK was much improved. I think it is a great tool that can support the teaching and learning of mathematics, especially in our situations where class sizes are too large and the support for teaching is non-existent."* (Dr. Michael Oyiengo, Maseno University)

“The academic year 2022/23 marked a significant turning point for MMUST in terms of delivery of mathematics courses. The Department of Mathematics was able to integrate STACK in the delivery of 8 courses. The implementation challenges notwithstanding, the support of interns in authoring the questions was incredible. The presentations by MMUST lecturers at the just-ended first African STACK conference was very revealing in terms of the impact of STACK on the teaching and learning experience. It was evidently clear that the integration of STACK in course delivery is positively correlated with student performance. We can't thank IDEMS enough for the support offered so far. We plan to upscale the integration of STACK in our courses in the coming academic year. In this regard, we are counting on the continued support from the interns.” (Prof. George Lawi, MMUST)

"Our journey with STACK began after attending a transformative workshop at Maseno University in 2019. Witnessing our students' remarkable progress using IDEMS-developed assessments for the 'Linear Algebra' course inspired us to embrace STACK in our teaching and assessments. We have since integrated STACK into our education system, offering personalized and immediate feedback, enhancing self-paced learning, and empowering our students to succeed. Our fruitful collaboration with IDEMS led to the successful implementation of a STACK-based Remedial Program, providing crucial support to a large number of students. As we continue our partnership, we are developing an online Mathematics support course for Natural and Social Sciences. The vast number of first-year students enrolled in this course spread across different campuses, demands innovative solutions, and STACK enables us to deliver timely guidance and support. Looking ahead, we are eager to leverage STACK's potential for the national exit exam, a crucial requirement for our final-year undergraduate students. We aim to provide invaluable preparation tools for this significant exam. Together with IDEMS, we are paving the way for a brighter future in mathematics education at Bahir Dar University." (Mebratu Fenta Wakeni, Bahir Dar University)

6. Challenges Encountered

The implementation of Open Stack Question Banks in African universities, while transformative, has not been without its challenges. This section discusses the key hurdles faced during the adoption and integration of Open STACK Question Banks across educational institutions in the continent.

1. **Training and Capacity Building:** While training was provided to mathematics graduates, ongoing support and capacity building are crucial for sustainable implementation. Ensuring that educators and technical staff are proficient in using the STACK system requires continuous training efforts, which can be resource-intensive and time-consuming.
2. **Adoption and Integration:** Encouraging widespread adoption of OQBs requires overcoming institutional resistance and fostering a culture of innovation in teaching and assessment. Educators may require convincing of the benefits of digital assessment tools over traditional methods, necessitating advocacy and demonstration of the effectiveness of OQBs in enhancing learning outcomes.
3. **Content Customization:** Tailoring OQBs to meet the specific educational contexts and curriculum requirements of diverse institutions across Africa poses a challenge. Adapting content to align with different educational standards and linguistic preferences

while maintaining quality and relevance requires careful planning and collaboration with local stakeholders.

4. **Curriculum Alignment:** Ensuring alignment between STACK question banks and existing curricula across diverse educational settings proved to be a complex task. Tailoring adaptive assessment tools to reflect regional educational standards and learning objectives required iterative adjustments and feedback from educators and subject matter experts. Continuous curriculum alignment remains essential to maximize the relevance and impact of STACK in supporting student learning outcomes.

5. **Resource Constraints:** Resource constraints posed significant challenges during the development and deployment of Open STACK Question Banks. Limited funding for technological infrastructure, including server maintenance and software updates, impacted the scalability and sustainability of the initiative. Collaborative efforts with funding agencies and strategic partnerships were essential in securing resources needed to expand access and enhance platform functionality.

Strategic Insights and Recommendations

Building on the successes and lessons learned from the implementation of Open STACK Question Banks, this section offers strategic insights and recommendations for sustaining and expanding this transformative initiative across African universities.

7. Next Steps:

1. Continuous Improvement and Expansion of Content:
   - Regularly update and expand the Open Stack Question Banks to cover a broader range of topics and levels within mathematics disciplines.
   - Incorporate feedback from educators and students to enhance the relevance and quality of questions, ensuring alignment with evolving educational standards and curriculum requirements.
2. Capacity Building and Training:
   - Provide ongoing training and professional development opportunities for educators and technical staff on using the STACK system effectively.
   - Expand training efforts to reach more institutions and ensure sustainability by empowering local trainers and educators to become proficient in authoring and using questions in the Open STACK Questions Banks.
3. Enhanced Integration and Adoption:
   - Foster greater adoption of Open STACK Questions Banks across additional universities and educational institutions by showcasing successful case studies and demonstrating the benefits of digital assessment tools.
4. Technical Support and Infrastructure Development:
   - Provide technical assistance and troubleshooting resources like servers to ensure smooth implementation and usage of Open STACK Questions Banks across diverse settings.

5. Community Engagement and Collaboration:
    - Strengthen partnerships with local and international stakeholders, including universities, educational organizations, and technology providers, to foster a collaborative ecosystem for developing and maintaining Open STACK Questions Banks.
    - Establish communities of practice and forums for educators to share best practices, exchange ideas, and collaborate on the development of new question sets and educational resources.
6. Evaluation and Impact Assessment:
    - Conduct regular evaluations and assessments to measure the impact of Open STACK Questions Banks on student learning outcomes, teaching effectiveness, and institutional capacity building.
    - Use data-driven insights to identify areas for improvement and inform strategic decisions for scaling and sustaining Open STACK Questions Banks across diverse educational contexts.

Bibliography

Emmaculate, O.,& Mary, S. (2024). *Building Open STACK Question Banks*. INNODEMS. https://shorturl.at/hYWuC

Focus: Other topics related to STACK.

Article number: 03

Data Procession Made Easy: A Python Tool for Extracting Information from Student Responses to STACK Questions

Jonas Lache*

Hochschule Ruhr West and Ruhr-Universität Bochum, Germany

Abstract

This paper presents the STACK Response File Processor, a newly developed tool for processing student responses to STACK questions. The tool, written in Python, consists of a user-friendly graphical user interface and facilitates the import of a Moodle quiz report in CSV format, giving the user a number of options such as the selection of input fields and potential response trees. The result is a CSV file that contains extracted information, including students' scores in specific subtasks. Aimed at assisting educators and practitioners in the efficient analysis of student responses to STACK questions, this paper introduces the recently developed tool and provides insight into its potential to streamline the processing of student response files provided by Moodle.

* Corresponding author: jonas.lache@hs-ruhrwest.de

1. Introduction

Learning management systems such as Moodle have become an integral part of university teaching. As a significant proportion of student learning takes place within these systems, a vast amount of data is generated, which can be utilised for a variety of purposes. For instance, Poellhuber et al. (2023, p. 592) use Moodle data for learning analytics with the objective to “help teachers identify students needing support, and to predict and prevent dropout” – and they are not the only authors with such an approach (e.g. Quinn & Gray, 2020; Raga & Raga, 2017). When mathematics questions created with the Moodle question type STACK are used in a course (see Sangwin, 2015), Moodle stores some data explicitly referring to the students' work on the tasks. For example, Wang et al. (2023) use data specific to STACK to create knowledge maps that display students' learning progress. In addition, Landenfeld et al. (2021) have developed a tool for the analysis of students' responses to a STACK question using a graphical approach.

This paper presents a new tool to facilitate the use of Moodle data on STACK questions. The user-friendly STACK Response File Processor extracts information from so-called Responses files that can be downloaded from a Moodle quiz after students have completed STACK tasks. First, these files are briefly introduced. After outlining the properties and features of the Response File Processor as well as its case of application, the text concludes with an outlook to possible future developments. This paper is a continuation of the work by Lache (2023).

2. The Moodle Responses output file

There are several ways to access data containing information about student activity in Moodle. However, in the context of STACK questions, the options vary in how useful they are. For example, while Moodle logs are suitable for learning analytics because almost all student activity is logged (see Rotelli & Monreale, 2023), these files do not contain information about individual student responses. A possibility to get data more specific to STACK is the Basic Question Use Report which is generated directly by the STACK plugin (see https://docs.stack-assessment.org/en/Authoring/Reporting). The report provides information about the actual answers given by students and is helpful for improving STACK questions (e.g. by adding more nodes to potential response trees to catch common mistakes). However, it only provides information at the level of STACK questions and not at the level of students and their individual responses. To get exactly that – a report of all individual student attempts, including their responses, which nodes in the potential response threes they traversed, and the score they received – the Responses output files are a sufficient choice (Lache, 2023).

These files can be accessed and downloaded (e.g. in CSV format) from the Moodle quiz. A Responses file is a table where each row represents a student's attempt at a STACK question. If the adaptive mode of the Moodle quiz is used (i.e. students can answer multiple times via clicking on a “Check” button until they submit the quiz), each row essentially represents one click on the “Check” button. The columns of the table contain different information on the attempts in each row (e.g. the time students spent on the quiz). An example of a pseudonymised Responses table is shown in Figure 1.

	A	B	C	D	E	F	
1	Alias	State	Started on	Completed	Time taken	Grade/10.00	Response 1
2	BO7JMWWmuX43Y8zi		5 April 2023 8:49 AM	5 April 2023 8:53 AM	4 mins 4 secs		Seed: 1646667833;
3	BO7JMWWmuX43Y8zi		5 April 2023 8:49 AM	5 April 2023 8:53 AM	4 mins 4 secs		Seed: 1646667833;
4	BO7JMWWmuX43Y8zi		5 April 2023 8:49 AM	5 April 2023 8:53 AM	4 mins 4 secs		Seed: 1646667833;
5	BO7JMWWmuX43Y8zi		5 April 2023 8:49 AM	5 April 2023 8:53 AM	4 mins 4 secs		Seed: 1646667833;
6	BO7JMWWmuX43Y8zi		5 April 2023 8:49 AM	5 April 2023 8:53 AM	4 mins 4 secs		Seed: 1646667833;
7	BO7JMWWmuX43Y8zi		5 April 2023 8:49 AM	5 April 2023 8:53 AM	4 mins 4 secs		Seed: 1646667833;
8	BO7JMWWmuX43Y8zi		5 April 2023 8:49 AM	5 April 2023 8:53 AM	4 mins 4 secs		Seed: 1646667833;
9	BO7JMWWmuX43Y8zi	Finished	5 April 2023 8:49 AM	5 April 2023 8:53 AM	4 mins 4 secs	10.00	Seed: 1646667833;
10	yLtYtNN3MJBNnkpz	Finished	5 April 2023 8:53 AM	5 April 2023 8:57 AM	4 mins 18 secs	0.00	Seed: 1579016987;
11							

Figure 1: Example of a pseudonymised Responses output file.

The “Response 1” column (or “Response 2” and so on if there are multiple questions in the Moodle quiz) contains a wealth of information about the students' attempts, encoded in strings. The following string is an example of an entry:

```
Seed: 1579016987; ans1: 2*x-4 [score]; ans2: 0 [score]; prt1: # = 1
| prt1-1-T; prt2: # = 0 | prt2-1-F
```

The sample string contains the value of the input fields (`ans1` and `ans2`) as well as the score achieved in each of the potential response trees (`prt1` and `prt2`) and whether they became active. These strings are human readable, but it is clear that it is much more effective to have a computer read the data. One way of extracting information from Responses files (especially the Response columns) is to use the Python programming language. Using libraries for data analysis, data manipulation and regular expressions, it is possible to extract information such as the value of input fields and the score achieved in potential response trees. To store this information, it is sensible to add a new column to the Responses table for each piece of extracted information. For more details on this approach and examples of Python functions, see Lache (2023). To make the approach easy and convenient to use for a wide range of STACK users, the free STACK Response File Processor has been developed.

3. The STACK Response File Processor

The STACK Response File Processor is a free and open-source tool written in Python. It has an intuitive and user-friendly graphical user interface implemented using the Python library `tkinter`. This section describes the features of the tool and how to use it. A summary is given in Figure 2.

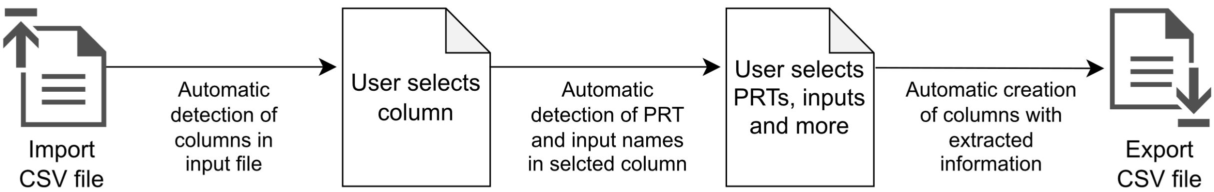


Figure 2: Steps when using the STACK Response File Processor.

After starting the Response File Processor, a button with the title “Open CSV file” appears (see Figure 3, left). Clicking on this button opens a file dialogue box asking the user to import a Responses file in CSV format (see Figure 3, right). It allows to browse for a file in the folder structure on the user's computer and then select it for import.

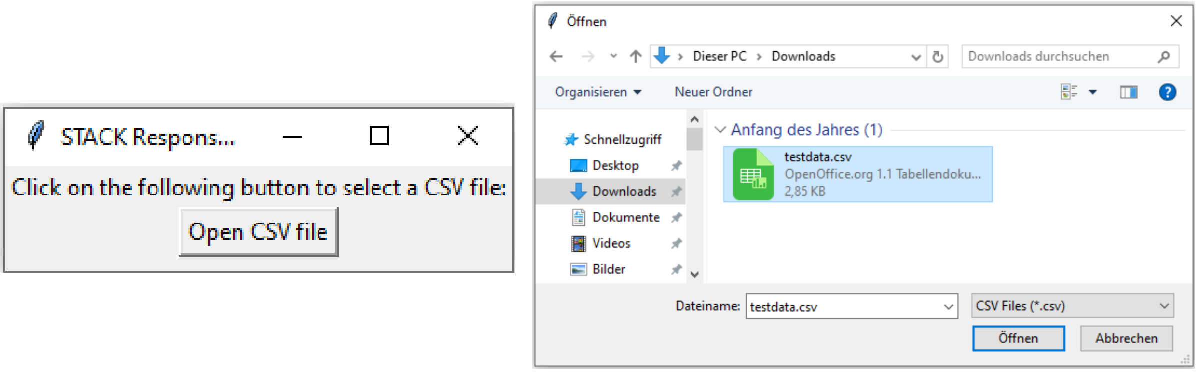


Figure 3: File dialogue when starting the tool and clicking on “Open CSV file”.

Once the import file has been loaded, the tool imports the data, processes it and lists all the columns present in the table. The user is then prompted to select a Response column, which refers to a STACK question and contains the information to be extracted. Once the user has selected a Response column, the tool automatically detects the names of the input fields and potential response trees present in the strings that are contained in the selected column. The Response File Processor then prompts the user to select the input fields and potential response trees for which columns are desired (see Figure 4). For each of the selected input fields, the tool creates two columns. The first one contains the value of the corresponding input field (e.g. a number, a term like `x^2+2*x` or a list like `[1,2,3]` that the students have entered). The second column contains the status of the input field, i.e. whether the students' answers were valid or not: in the case the students received a score for a valid answer, the value here is “score”. If the answer is considered valid by STACK but the students did not yet receive a score, the value is “valid”. And if the answer is considered invalid, the value in the column is “invalid”. Two columns are also created for each of the potential response trees that the user selected. The first column contains the information whether the response tree was active (True or False). The second column contains the score that the students achieved in the given response tree (a number between 0 and 1).

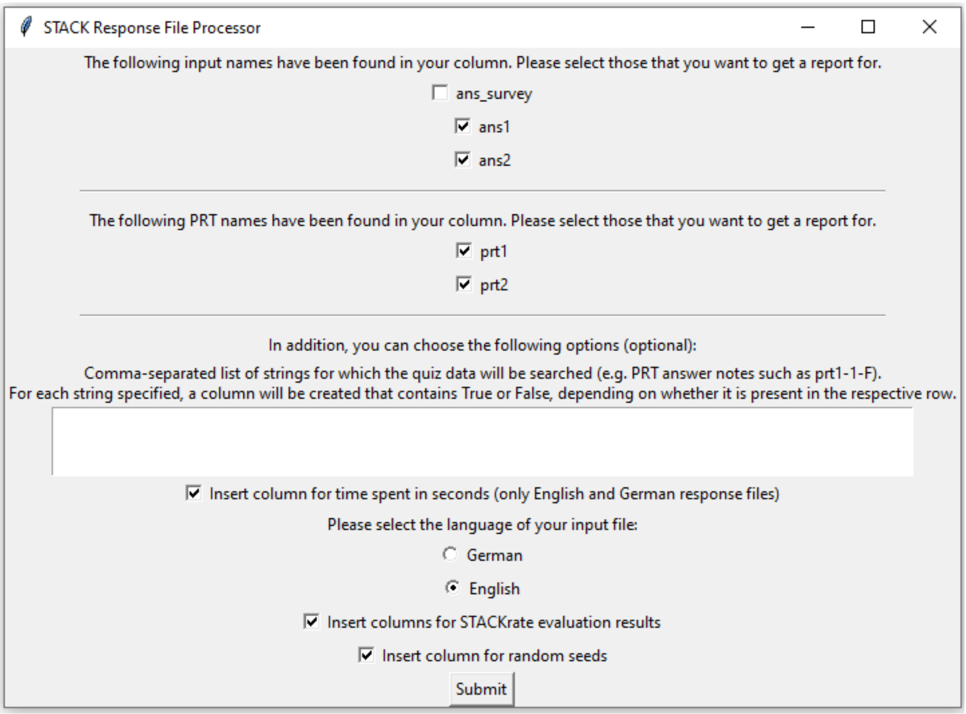


Figure 4: Selection options for the user.

In addition to the choice of input fields and potential response trees, users can select a number of optional settings. Selecting the "Insert column for time spent in seconds" checkbox will create a column titled "Seconds spent". This column is a conversion of the "Time spent" column as described by Lache (2023): The "Time spent" column is included by default in each Responses file and contains information about how long students spent on the Moodle quiz. However, as it does not contain numbers but strings such as "4 mins 18 secs" which are not easily parsed, it is sensible to convert these values to seconds. The Response File Processor is able to do this if the Responses file has been exported from English or German Moodle systems. When selecting the checkbox mentioned above, the user must choose whether a German or an English input file is being used (see Figure 4).

Another optional setting is the checkbox "Insert columns for STACKrate evaluation results" (see Figure 4). STACKrate is a free and open-source tool that makes it easy for teachers to implement an evaluation of STACK questions. It allows students to answer evaluation questions using star ratings and an open text field. The rating results are stored in a hidden STACK input field and are therefore part of the Response columns in Moodle Responses output files (see Lache & Meißner, 2022; Meißner & Lache, 2024). If the checkbox is checked, the Response File Processor will extract and process the STACKrate evaluation results that are stored in the underlying data. The results are then added to a column containing all STACKrate ratings as lists (e.g. `[5,4,5]`, where each entry represents the number of stars a student gave for a particular evaluation question). If available, the tool will also create a column with the free comments written by the students.

In version v0.2 of the Response File Processor, two more options have been added. Firstly, the tool can now create columns with a True or False value depending on whether strings specified by the user are contained in the Response column in the corresponding row. A free text field can be used to enter, for example, answer notes (e.g. `ATList_wrongentries` or `prt1-1-F`). This option is useful if users are interested in whether students made a particular mistake or received particular feedback, and makes the Response File Processor more adaptive and flexible. By ticking the also newly added "Insert column for random seeds" checkbox, it is now possible to have a column containing the random seed used in each student attempt (see Figure 4).

Once all the options have been selected as desired, a "Submit" button must be clicked. On the next page, clicking a "Save CSV file" button opens a file dialogue similar to the one shown in Figure 2 (right). The user is prompted to browse to a desired folder on the computer and enter a filename for the export file. The Response File Processor will then save the output file in CSV format. The file is now ready for analysis using a spreadsheet or statistics program.

4. Outlook

The aim of the STACK Response File Processor is to facilitate the processing of Moodle Responses files that contain information about students' work on STACK questions. By using a graphical user interface, the approaches and Python functions presented by Lache (2023) became accessible to all STACK users – even those with little experience of Python. Responses files processed with the tool can be used for a variety of analyses, such as statistics on the scores achieved by students on particular subtasks of a STACK question, their time spent on the Moodle quiz, and the frequency with which different nodes of potential response trees were traversed by students.

There are several ideas for improving the tool, which still is in an early stage of development. At the moment, if the Moodle quiz to which the Responses file refers has multiple STACK questions, users have to run the Response File Processor multiple times and select the

Response columns one at a time. Adding a feature that allows users to select multiple columns from the input file at once would make this process easier. Another potentially helpful feature would be to allow users to select the columns to be created for each selected input field and potential response tree. Currently, the same columns (two for each input field and response tree, see section 3) are created by default for each selected input field and response tree and there is no option to only select the desired columns. Adding such a feature could make the tool better suited to the needs of users who want to have maximum control over the columns generated. Another idea would be to have the Response File Processor automatically calculate statistics, such as the average score achieved by students in a particular potential response tree of a STACK question. Similarly, graphics based on the processed data, such as charts visualising the above-mentioned statistics, could be automatically generated. If desired, the statistics and graphs could be automatically exported along with the data, which is exported as a CSV file anyway. Finally, supporting more languages (and not just English and German) for converting the strings in the “Time spent” column to seconds would make this feature available to a wider group of STACK users.

The project is hosted on GitHub (see https://github.com/jonaslache/STACK-Response-File-Processor) where the source code and documentation can be found. Anyone can use the tool for free, regardless of whether they use Windows, macOS or a Linux distribution as their operating system. Bug reports, feature requests and community contributions are welcome.

Bibliography

Lache, J. (2023). Processing Student Responses to STACK Questions in Moodle Quizzes. *Contributions to the International Meeting of the STACK Community 2023*. https://doi.org/10.5281/ZENODO.8147587

Lache, J., & Meißner, D. (2022). How to STACKrate: The Development of an Intuitive Evaluation Tool. *International Journal of Emerging Technologies in Learning (iJET)*, *17*(23), 7–14. https://doi.org/10.3991/ijet.v17i23.36527

Meißner, D., & Lache, J. (2024). New Developments Around the Evaluation Tool STACKrate. *Proceedings of International Meeting of the STACK Community 2024.* 75-80.

Landenfeld, K., Eckhoff, M., & Priebe, J. (2021). Graphical Visualization of STACK Response Analysis. *Contributions to the International Meeting of the STACK Community 2021*. https://doi.org/10.5281/zenodo.4915905

Poellhuber, L.-V., Poellhuber, B., Desmarais, M., Leger, C., Roy, N., & Manh-Chien Vu, M. (2023). Cluster-Based Performance of Student Dropout Prediction as a Solution for Large Scale Models in a Moodle LMS. *LAK23: 13th International Learning Analytics and Knowledge Conference*, 592–598. https://doi.org/10.1145/3576050.3576146

Quinn, R. J., & Gray, G. (2020). Prediction of student academic performance using Moodle data from a Further Education setting. *Irish Journal of Technology Enhanced Learning*, *5*(1). https://doi.org/10.22554/ijtel.v5i1.57

Raga, R., & Raga, J. (2017). Monitoring Class Activity and Predicting Student Performance Using Moodle Action Log Data. *International Journal of Computing Sciences Research*, *1*(3), 1–16. https://doi.org/10.25147/ijcsr.2017.001.1.09

Rotelli, D., & Monreale, A. (2023). Processing and Understanding Moodle Log Data and Their Temporal Dimension. *Journal of Learning Analytics*, *10*(2), 126–141. https://doi.org/10.18608/jla.2023.7867

Sangwin, C. (2015). Computer Aided Assessment of Mathematics Using STACK. In *Selected Regular Lectures from the 12th International Congress on Mathematical Education* (pp. 695–713). Springer International Publishing. https://doi.org/10.1007/978-3-319-17187-6_39

Wang, L., Nakamura, Y., & Nakahara, T. (2023). Visualization of Knowledge Map based on STACK Answer Data. *Contributions to the International Meeting of the STACK Community 2023*. https://doi.org/10.5281/ZENODO.8147592

Information on the author

Jonas Lache

He is a research assistant at the Ruhr-Universität Bochum and a PhD student in mathematics education. He is also a member of the e-learning team at the Hochschule Ruhr West, where he is responsible for the Moodle question type STACK.

Focuses: JSXGraphs, interactive STACK exercises, ALepa, learning sequences.

Article number: 04

Designing graphical physics problems with JSXGraph

Michael Kubocz*, David Lauter, Stefan Roth

RWTH Aachen University, Department of Physics, Germany

Abstract

Learning management systems, such as Moodle, offer the possibility to embed supplementary material to any course, which can be utilized not only for a summative and formative assessment, but also for student self-study. This includes single/multiple choice questions, problems that require algebraic and/or numerical answers, with or without units for the latter. The STACK plugin is a powerful tool for creating such digital learning material in order to test students' understanding and knowledge of the topic. In addition, using the open-source JavaScript library JSXGraph, the STACK plugin also allows the integration of graphical elements, such as geometric constructions, interactive plots, 2D and 3D visualization of objects, and even animations. The graphical content can be linked to randomized parameters of a given problem itself, thereby augmenting the specific problem individually. In addition, a new type of problems can be implemented which students have to solve graphically. In this case the students can provide an answer by moving or modifying given graphical objects. In this presentation we demonstrate some examples of STACK problems employing JSXGraph, that can be used within a university physics course.

* Corresponding author: kubocz@physik.rwth-aachen.de

1. Introduction

In the STEM field, there exist ample initiatives to provide mathematical exercises for prospective and current students. Although the available learning management systems (ILIAS, moodle) allow the use of such exercises, designing and formulating of the questions and their answers in an error-free, comprehensive, and interactive way remain very time-consuming for teachers of physics and technology subjects. The aim of the ALepa (Adaptable learning sequences for basic physical and technical training, or in German: **A**daptierbare **Le**rnsequenzen für die **p**hysikalisch-technische Grund**la**genausbildung) project is to develop interdisciplinary usable learning sequences based on structured task collections for the physical-technical basic education in the study programs physics, electrical engineering, and related courses, and publish on ORCA.nrw. It is a collaboration of six universities, namely, Dortmund University of Applied Sciences and Arts (FHDo), Cologne University of Applied Sciences and Arts (THK), Hamm-Lippstadt University of Applied Sciences and Arts (HSHL), Aachen University of Applied Sciences and Arts (FHAc), Ruhr University Bochum (RUB), and RWTH Aachen University.

The learning sequences are based on the principles of the Cognitive Load Theory (Sweller, 1988, 2006; van Gog, Paas & Sweller, 2010), namely, that worked-out examples have progressively less help (e.g. hints, self-explanations) and that the level of difficulty of exercises increase gradually up to including problem solving tasks in order to meet an expert-reversal effect (Renkl, 2005, Kalyuga, 2007). Each learning sequence has a thematic focus, e.g. mechanics, acoustics, or optics, which makes it possible to easily embed a selected learning material within that topic into the corresponding course. This approach is in accordance with Bauer's recommendation (Bauer et. al., 2019), according to which learners with little prior domain-specific knowledge need guided sequencing and a balanced division between text-only and interactive contents. While prompts and feedback on (self-test) tasks provide the learners with hints and links for further studies (especially for case studies), tasks of varying levels of difficulty help to compensate for their heterogeneous prior knowledge and learning speeds. By using the direct feedback from the automated (self-)tests, students can better organize their own learning pace, and effectively control the increase in the level of difficulty by themselves befitting their pace. This approach enables lower-performing students to also experience competence, which helps to uphold motivation (Deci & Ryan, 1993). Though the current learning sequences are designed in the aforementioned way, we would like to venture one step further, and replace, where possible, non-interactive contents with interactive graphs with the help of the JSXGraph library (Gerhäuser et al, 2010). In these proceedings, we will mainly focus on STACK exercises concerning interactive graphs, and divide them into two categories. In the first category the students are to construct the solution via manipulation of one or more elements of the graph, such as movable vectors, points, or certain portions of the graph. In the second category the interactive graph with movable parts like vectors, points etc. is meant to serve only as an additional help to enhance the overall comprehension of the given problem to successfully construct the corresponding algebraic or/and numerical solution.

2. Examples

The following example (Figure 1) of a STACK exercise of the first category deals with the decomposition of the gravitational force $-\vec{F}_G$ (green vector) pulling on a weight (grey triangle) attached by two ropes of negligible masses (dashed lines) to two facing walls. The absolute value of $-\vec{F}_G$ and the angles of the two ropes with respect to the horizontal (red) are randomly chosen for this exercise. The calculation of the solutions, which, in this case are the coordinates of the forces $\vec{F}_L$ and $\vec{F}_R$ along the ropes (blue vectors), are done completely within the STACK environment. These coordinates are then passed to the embedded JSXGraph environment, which ensures a better adjustment of the exercise, if needed. To solve this exercise, students are supposed to bring the blue vectors to their correct locations, either by guessing/estimating, or calculating the corresponding coordinates while also taking the measure of the units of length involved into account. When the solution is submitted, the sample solution (red and black vectors) is shown together with that of the student's (blue vectors) (Figure 2). An algebraic master solution also follows, if available and/or relevant. Furthermore, a sufficiently large acceptance (error) interval around the actual solution (i.e. numerically calculated coordinates) is allowed in the presentation of the graphical solution to avoid problems with the input. In this exercise a relative error interval of 10% around the correct solution was used. The difficulty level of this exercise can be boosted, if needed, simply by removing the fixed initial points of the blue vectors.

The code of the last exercise can also be partly recycled for a new type of exercise, which, on the one hand, has an old-style numerical solution as its answer, while, on the other hand, is additionally accompanied by an interactive picture to enhance the comprehension of the work required to be done (Figure 3).

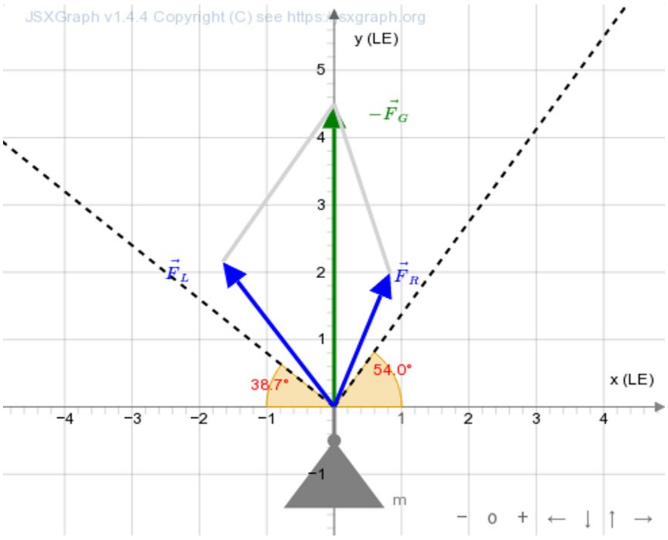


Fig. 1: Decomposition of the force -F ⃗_G into the two forces F ⃗_L and F ⃗_R (blue vectors) along the dashed lines.

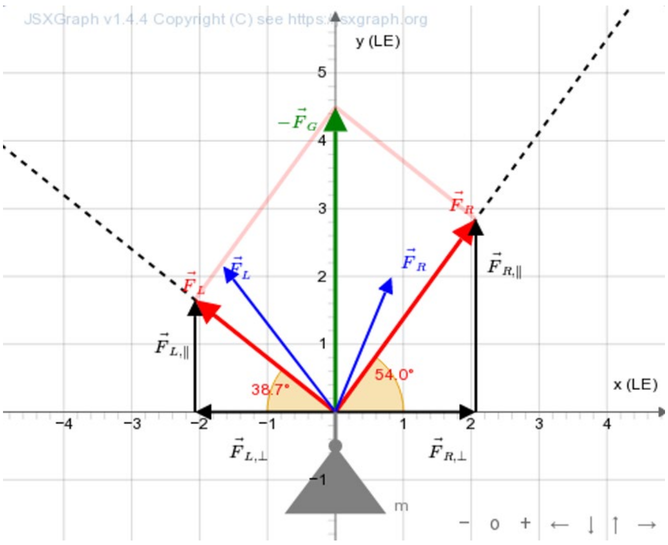


Fig. 2: Sample solution together with student's solution of the previous exercise.

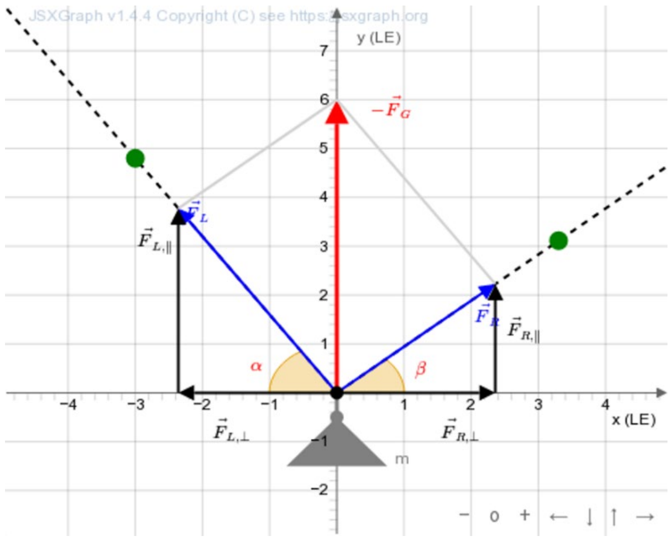


Fig. 3: Interactive picture in which the given force -F ⃗_G for any angle α and β can be decomposed along the two dashed lines by moving the two green dots

In this case students can arbitrarily move both green dots which change both angles α and β, and show the corresponding change in the decomposition of the force $-\vec{F}_G$ along the dashed lines with the help of the blue and black vectors. This visualisation is meant to not only improve the understanding of the exercise, but also to aid finding the corresponding ansatz to proceed with the numerical solution of such an exercise with random input parameters.

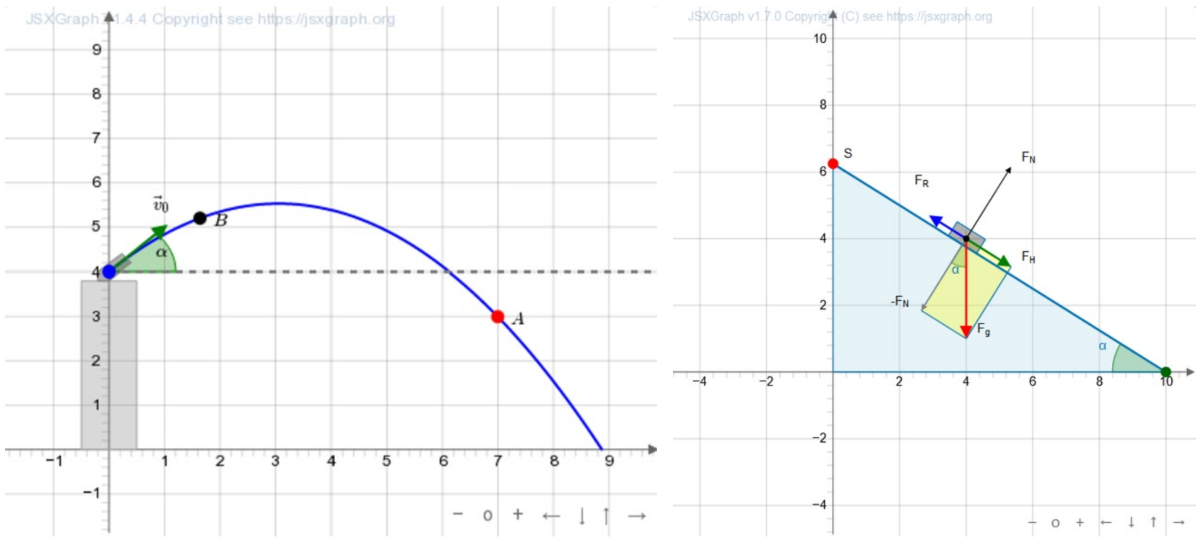


Fig. 4: Determination of the maximal potential and kinetic energy on a trajectory of a projectile via movement of dots A and B (left panel). Determination of the minimal angle of the plane to overcome the friction force via adjustment of the dot S (right panel

Two further examples of exercises are shown in Figure 4. The left panel of Figure 4 depicts the trajectory of a projectile (parabola) under gravitation. In this exercise students are supposed to pinpoint the locations of the maxima of the potential and kinetic energies by sliding the dots A and B to their corresponding locations. The right panel of Figure 4 shows a rectangular object placed on a tilted plane, where the exercise is to determine the maximal angle at which this object overcomes the frictional force and starts to slide down the plane under the action of gravity. By adjusting the angle alpha with the dot S the forces $\vec{F}_H$ (slope downforce), $\vec{F}_N$ (normal force) and $\vec{F}_R$ (friction force) can be changed, while $\vec{F}_g$ (weight force) remains unchanged. For the case of $|\vec{F}_H| = |\vec{F}_R|$, the corresponding angle can then be read off. Both exercises here aim at enhancing the comprehension of the corresponding topics.

3. Summary and Outlook

Our goal is to enrich existing physics problems in both physics and technology subjects by replacing non-interactive content with interactive graphs based on the JSXGraph library, which can both be used in stand-alone STACK exercises and in the context of creating as well as further developing learning sequences. The interactive graphs should not merely replace static pictures, but serve as additional visual assistance to successfully reach the solution of an exercise, while simultaneously enhancing the overall comprehension of the topic at hand. However, a second equally significant objective of this whole venture should also include a proper judgement of the impact of the same, both from the students' and the teachers' perspectives, and utilise their feedback to incorporate effective modifications as and where needed.

Bibliography

Bauer, M., Schuldt, J., Bau, B., Webers, M. J., Krömker, H. (2019). *Förderung der Lernmotivation durch adaptives E-Learning*, Lecture Notes in Informatics, 17. Fachtagung Bildungstechnologien, 163-174.

Deci, E. L., & Ryan, R. M. (1993). *Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik*. Zeitschrift für Pädagogik 39, (2), S. 223-238.

Gerhäuser, M., Valentin, B., Wassermann, A.: *JSXGraph – Dynamic Mathematics with JavaScript*, in The International Journal for Technology in Mathematics Education (IJTME), ISSN: 1744-2710, 17 4 (2010)

Kalyuga, S. (2007). *Expertise Reversal effect and its implications for learner-tailored instruction*. Educational Psychology Review (2007) 19:509-539.

Renkl, A. (2005). *The worked-out examples principle in multimedia learning*. In Mayer, R.E. (Ed.), The Cambridge Handbook of Multimedia Learning. Cambridge: Cambridge University Press.

Sweller, J. (1988). *Cognitive load during problem solving: Effects on learning*. Cognitive Science, 12, 257–285.

Sweller, J. (2006). *The worked example effect and human cognition*. Learning and Instruction, 16(2), 165–169.

van Gog, T., Paas, F. & Sweller, J. (2010). *Cognitive Load Theory: Advances in Research on Worked Examples, Animations, and Cognitive Load Measurement*. Educ Psychol Rev 22, 375–378. https://doi.org/10.1007/s10648-010-9145-4

Information on the authors

Dr. Michael Kubocz

has a PhD in theoretical particle physics, was a postdoc at two institutes at RWTH Aachen, and taught physics and maths for almost six years at both RWTH and DHBW Mannheim. He currently works at RWTH to develop e-exercises for two projects under OER.

David Lauter

is currently completing his Bachelor's thesis in experimental particle physics at RWTH Aachen and has been working as a student assistant in the field of scientific education for approximately two years.

Prof. Stefan Roth

is professor of physics at RWTH Aachen University. He is using e-learning in his classes since many years. He is co-author of the German-wide Online bridging course physics, which tries to bridge the gap between high school and university.

Focuses: Automation of question and feedback generation; STACK in teaching or exams.

Article number: 05

Enhancing Statistics Learning through Automated STACK Testing

Elena Safiulina, Natalja Maksimova*, Oksana Labanova

TTK University of Applied Sciences

Abstract

This article aims to explore the possibility of creating STACK tests with generated data and automating feedback mechanisms in the context of a statistics course at the TTK University of Applied Sciences (TTK UAS), thereby presenting a novel experience in utilizing STACK.

The challenge stems from the limited number of contact hours available to lecturers for an in-depth study of material on classical theoretical distributions of random variables, necessitating alternative methods to enable students to thoroughly study and reinforce the statistics curriculum and monitor the acquisition of their knowledge.

Keywords: Discrete Random Variable, Continuous Random Variable, Teaching and Learning of Statistics, e-Assessment, CAA.

* Corresponding author: natalja.maksimova@tktk.ee

1. Introduction

Distribution laws are crucial in statistics, as they help elucidate and interpret the behavior of random events and their probabilities. Random variables have become essential across almost all research fields, including physics, chemistry, engineering, and particularly in biological, social, and management sciences. These variables are assessed and examined based on their statistical and probabilistic characteristics, with the distribution function being their primary attribute (Forbes et al., 2011). These distributions are mathematical models that describe how the values of a variable are distributed or the probability of a certain event occurring. Many physical systems can be modeled by understanding that the data follows a specific distribution, allowing us to predict average outcomes. Researchers can draw conclusions about data that is slow or expensive to collect, or that cannot be collected for other reasons. Essentially, distributions are a fundamental component of statistical reasoning (Wild, 2006). Based on the laws of distribution, the population is estimated and hypotheses are tested.

Distribution laws fall between descriptive and summary statistics. Statisticians look at variation through a lens which is “distribution” (Wild, 2006). Unfortunately, it is not always possible to build a complete statistics course for students, especially for students of sessional training, where there are not enough classroom hours available to lecturers for in-depth study of material on classical theoretical distributions of random variables, which requires alternative methods that allow students to study and consolidate the curriculum on the subject of statistics, as well as monitor the acquisition of your own knowledge when completing the topic independently.

The practical task will be to describe a random variable. Regardless of the type of distribution law, the algorithm for solving problems is universal and consists of three stages that present a set of data through tables, graphs and numerical characteristics. This type of task can be organized with the help of STACK, because only this tool allows us to divide the question into related parts, creating several interrelated test questions (Safiulina et al., 2021).

The subject of statistics takes place in six different groups during one semester. Students learn in different learning formats using their Moodle courses. On average, a total of 120 students study. There are 6-8 different random size describing events in the homework, which is very few. It is necessary to reproduce these word problems by changing the probability of the occurrence of the event and the number of objects.

STACK is ideal for completing and expanding the number of variants (Safiulina & Labanova 2020), as it generates random task variants and automatically calculates answers, which is more efficient than conventional tests. The student solves tasks in MS Excel, R or other statistical software, and the STACK evaluation system is designed to check the student's solution and create a feedback evaluation schedule. The purpose of checking is to make sure that the learning tasks or goals have been completed correctly, and the teacher reports on the achievement through feedback. Homework is complete when the entire task is completed. If the student has a misunderstanding, the work shows it, and the lecturer has the opportunity to give hints already in the feedback of the work, because the assessment must help the student learn the subject better. The STACK question, with its interactive multiple-choice question mode, fits the bill perfectly.

STACK has been used effectively for a long time in various fields such as physics (Kröger & Schwarz, 2020; Schmitt & Spatz, 2020), economics (Riebe & Varmaz, 2021), graph theory ( ) and mathematics, including linear algebra (Härterich, 2019), analysis (Chongchitnan & Harrison, 2021), geometry (Safiulina, et al., 2023), and statistics (Hooper & Jones, 2023). In statistics, only calculation is applied at the level of a specific formula. There is not yet a solution to a complex task in practice: table, graphs, numerical characteristics.

In this paper, the authors have not yet worked with the creation part of the different branches of the evaluation graph. The focus of the article is aimed at creating tasks, the use of the

STACK system for creating a statistics task is illustrated using the example of discrete binomial distribution.

2. Methodology

The features of the STACK question and its creation will be illustrated by considering one of the homework problems that needs to be translated into an electronic test. The electronic test is planned to be introduced into TTK UAS in the 2023/24 academic year to assess all students taking this course.

> Task. "*A target is shot at independently 7 times. The probability of hitting the target with one shot is 0.6. Random variable X - the total number of hits out of 7 shots.*"

Regarding the description of a random discrete binomial value, in which the description depends on the parameters `n` number of trials and with the probability of success on a single trial denoted by `p`.

```
p:                          rand_with_step(0.3,0.9,0.05)
n:
                            rand_with_step(4,7,1)
```

With such parameters, 52 different options are generated only for this task, where using the function `rand_with_step(min,max,step)` the number of generated numbers is equal to $(max - min + 1)/step$ and, if necessary, even more options can be obtained.

Skills assessed in this e-assessment include the ability to work in MS Excel using the Moodle template provided in the course, where the student completes all necessary calculations.

Description using table

Table of distribution of random variable $X \sim B(n, p)$, where probability distribution probability of exactly k successes in n trials is given by

$$P(X = k) = C_n^k p^k q^{n-k},$$

consists of three lines, each of which is entered in accordance with a certain formula and is subsequently checked separately. To provide access to each row, they must be initialized as separate objects in the test (Fig. 1).

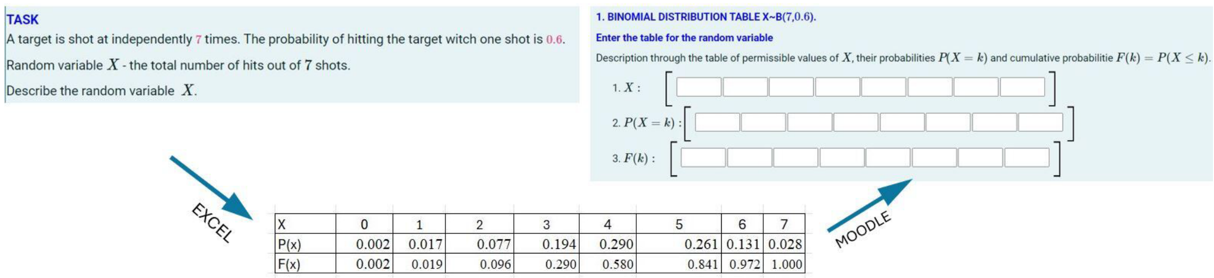


Figure 1. Table of distribution in STACK task

Incorrect calculation of probabilities at this stage can lead to errors in graphs and calculations at subsequent stages. Therefore, when students interact with a computer-aided assessment (CAA) system, it is very important to ensure that the calculations and tables are filled out correctly (Gill & Greenhow 2008).

Feedback is the point in the learning process when students receive the most personalized instruction. Effective feedback should be accessible and useful to students, improving their understanding of their studies and improving their future performance (Obilor, 2019). In STACK questions, specific feedback is determined through a potential response tree (PRT). A PRT may include one or more nodes, with each node potentially having multiple child nodes.

According to the PRT algorithm, student responses are initially evaluated against a property set in the root node. Based on whether this evaluation passes or fails, the process continues in the relevant child node. This recursive algorithm proceeds until no further child nodes exist. During this process, partial scores can be added or subtracted at each node along the evaluation path. Additionally, feedback messages can be generated at each node using information from the student's response. Although multiple PRTs can be defined for a single question, there is only one section for general feedback, ensuring that all students receive the same general feedback regardless of their responses (Knaut et al., 2022).

For a task under development, feedback is used with several attempts or an answer before correction, which allows the student to independently find the correct answer, offering the possibility of one or more attempts to solve it (Shute, 2008). This type of feedback is suitable for testing a group of students to improve higher order learning that involve critical thinking, analysis, synthesis, and evaluation (Mertens et al., 2022). In the Moodle test tool, the "Interactive with multiple tries" mode was used, which allows you to immediately check the entered answers and, in case of an error, provide the student with a hint that is written in the STACK system (Fig. 2a illustrates the passage of each test question) and the opportunity to start a new attempt with the initially generated task parameters. In case of incorrect input, the Student is offered three levels of type of hints and a general standard feedback for incorrect answers (Standard feedback for incorrect input), reminding about the rules for entering answers (Fig. 2b).

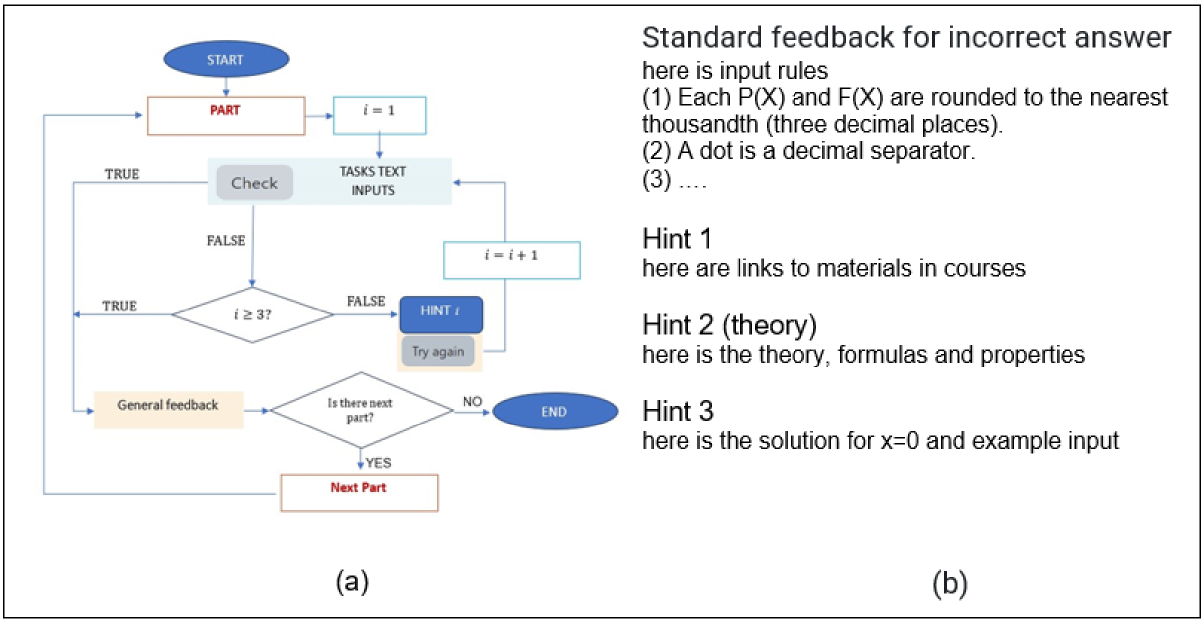
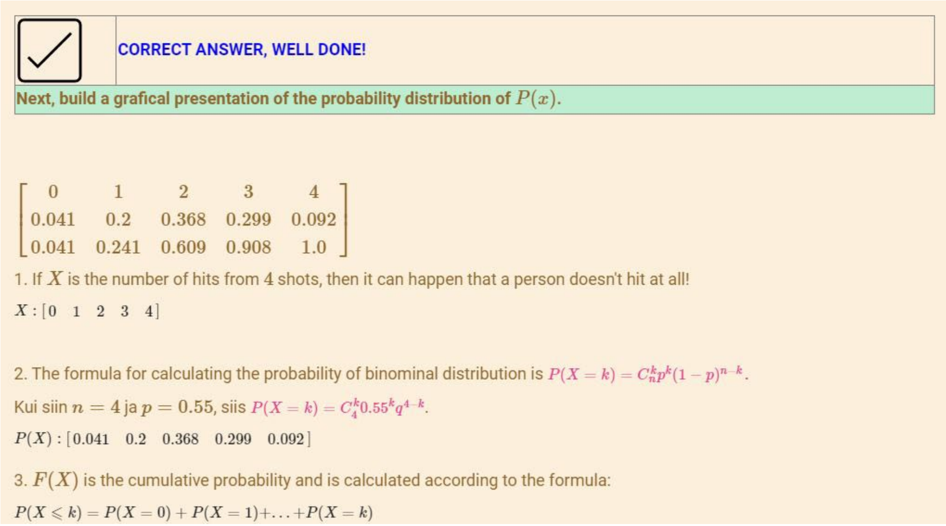


Figure 2: Structure of standard feedback for incorrect answers

Two hints provide links to educational videos, formulas, properties, and only the last hint provides a calculation for the first value with a full explanation of the formulas used, but does not give the correct results, which allows the student to be activated for independent work with the topic material. If the answer is correct, the correct answers are also indicated (Fig. 3).


CORRECT ANSWER, WELL DONE!

Next, build a grafical presentation of the probability distribution of $P(x)$.

$$\begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0.041 & 0.2 & 0.368 & 0.299 & 0.092 \\ 0.041 & 0.241 & 0.609 & 0.908 & 1.0 \end{bmatrix}$$

1. If X is the number of hits from 4 shots, then it can happen that a person doesn't hit at all!

$X : [0 \quad 1 \quad 2 \quad 3 \quad 4]$

2. The formula for calculating the probability of binominal distribution is $P(X = k) = C_n^k p^k (1-p)^{n-k}$.

Kui siin $n = 4$ ja $p = 0.55$, siis $P(X = k) = C_4^k 0.55^k q^{4-k}$.

$P(X) : [0.041 \quad 0.2 \quad 0.368 \quad 0.299 \quad 0.092]$

3. $F(X)$ is the cumulative probability and is calculated according to the formula:

$P(X \leqslant k) = P(X = 0) + P(X = 1) + \ldots + P(X = k)$

Figure 3: Feedback on the correct answer]

This approach is applied to all parts of the task and allows the student to arrive at correct calculations and form new knowledge and build confidence in learning, which is very important when completing the topic independently.

Graphical presentation

The second part of the question consists of plotting two graphs: a graphical representation of the probability density function (PDF) and the cumulative distribution function (CDF) and has therefore been split into two items.

The Probability Density Function (PDF)

To answer the question about the graphical presentation of the Probability Density Function (PDF), students create a histogram with a distribution curve as a bar graph, which teaches them to use the more advanced features of MS Excel. The following graphs are often used in programs for displaying the PDF: a set of graphs, a set of lines, and a distribution curve (Fig. 4). The last one was chosen by the authors to be used in the task.

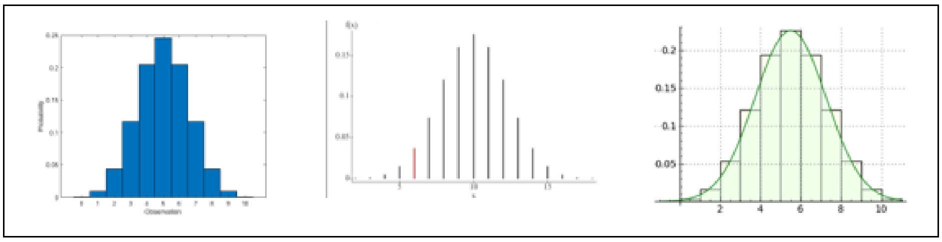


Figure 4: Graphs used in programs for displaying the PDF

For graphical representation in STACK questions, there are features such as the `plot()` function for basic graphs that have been selected, GeoGebra integration (Lutz, 2019) and JSXGraph for more advanced dynamic graphs with visual effects display (Oulu University Moodle, 2024). The STACK `plot()` command is defined as a wrapper for the Maxima command `plot2d()`. Not all `plot2d` functions are available through `plot()` (HFT Stuttgart Moodle, 2024).

The usual `plot()` functionality for STACK questions is designed to visualize a list of points, points connected by line segments, and curves of explicit and parametric functions (Department of Mathematics and Statistics, 2024). The histogram plot code with the distribution curve was created through a set of points connected by segments (lines), which was visualized along with the density function. When creating schedule options, it is important to reduce the likelihood of accidentally selecting a schedule without completing the task. Direct STACK questions do not support changing the order (shuffling) of input options in the radio answer type across questions and within each attempt. The random shuffling of options is recorded in the question code (Aalto OpenLearning, 2024). To avoid finding the correct answer by eliminating incorrect options, the proposed graphs differ slightly from the correct graph. Thus, the likelihood of choosing a schedule at random decreases with each attempt.

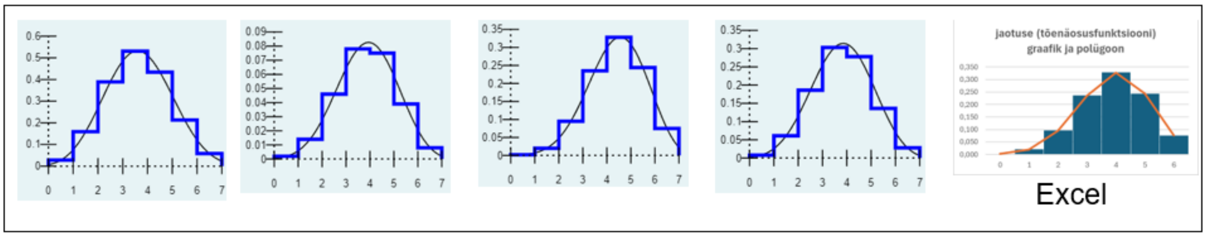
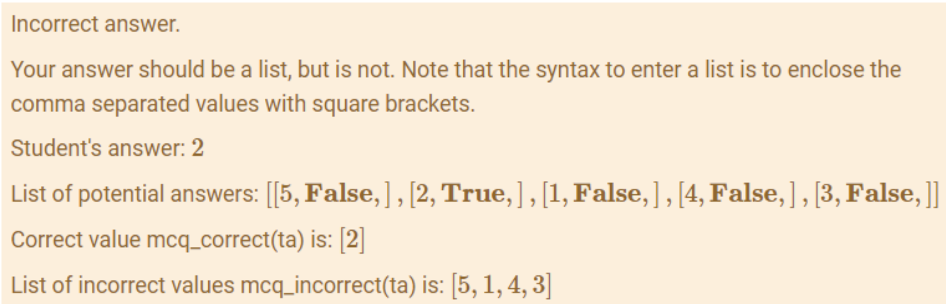


Figure 5: Proposed graphs

STACK questions use a comparison of two expressions to check the correctness of the answer: `Student Answer` (CAS expression, which is the student's input answer) and `Teacher Answer` (the calculated model answer). The model's response becomes calculations that are stored in code variables and specified in the model's response in the input block. This makes it possible to determine whether they satisfy certain mathematical criteria. By default, the check is aimed at establishing the algebraic equivalence of expressions `AlgEquiv` (Fig. 6).

Figure 6: Checking the answer

The student selects from a list of collections of answer choices, and the choice number is a number that is recorded as his answer (`SAns`). In this case, it cannot be directly compared with the correct response of the model (`TAns`), which is an element of the list due to differences in data types (Fig. 7).


Incorrect answer.

Your answer should be a list, but is not. Note that the syntax to enter a list is to enclose the comma separated values with square brackets.

Student's answer: 2

List of potential answers: $[[5, \mathbf{False},], [2, \mathbf{True},], [1, \mathbf{False},], [4, \mathbf{False},], [3, \mathbf{False},]]$

Correct value mcq_correct(ta) is: $[2]$

List of incorrect values mcq_incorrect(ta) is: $[5, 1, 4, 3]$

Figure 7: The student's correct choice does not coincide with the correct answer

It is necessary to solve this discrepancy in a separate code so that the algebraic equivalence of these two expressions is correctly established `SAns` and `TAns`.

The Cumulative Density Function (CDF)

The initial goal of the Cumulative Distribution Function (CDF) question was to know the type of CDF graph of a discrete random variable, which was taken into account when creating test selection options. Excel does not have a graph type for constructing the Cumulative Distribution Function (CDF) and in normal work it was not asked to create this graph, the emphasis was on the theoretical understanding of its step form (Fig. 8a).

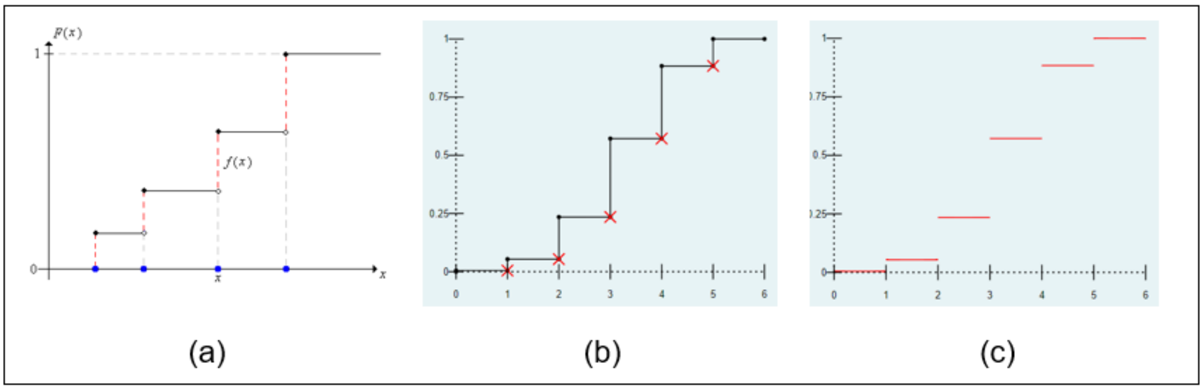


Figure 8: Graphical representation of CDF

The `plot()` function has the ability to change point style through thickness, color, and type, and line style through thickness, color (Edwin & Woollett, 2014). The absence of a dotted line type makes it impossible to create a similar graph and therefore the graph (Fig. 8b) is not correct. The second option, using a piecewise function, was the correct option for this task (Fig. 8c).

Numerical measures

The last part of the question checks the calculation of numerical characteristics: Numerical measures: Mean (EX), Dispersion (DX), Standard Deviation (σ), Mode (Mo), Median (Me), Lower Quartile (Q1), Upper Quartile (Q3).

When finding the first three numerical characteristics, formulas are used, and for the rest, a frequency table is used. Therefore, part of the question was divided into two points (Fig. 9).

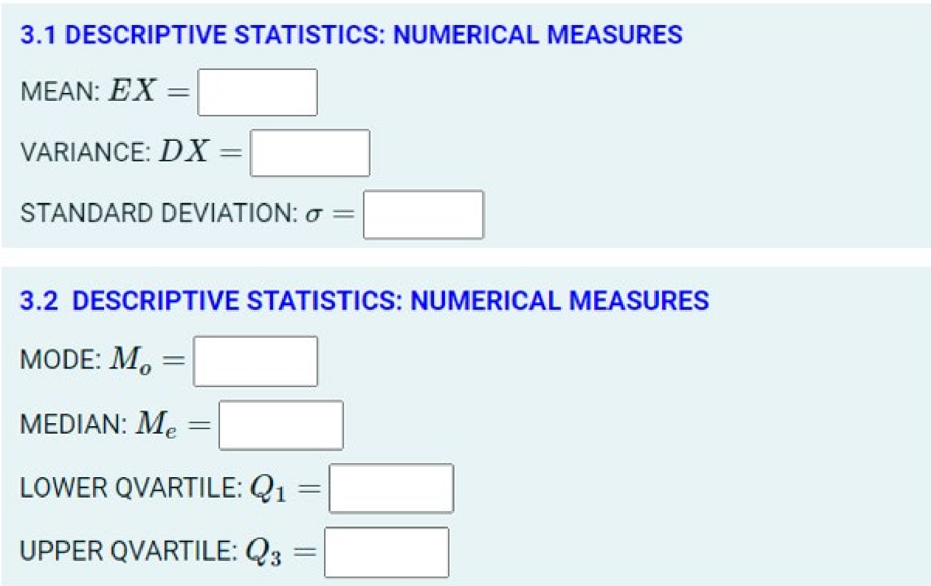


Figure 9: Numerical measures

Dividing this part into two questions simplifies the process of review, feedback, and analysis. It helps students focus on different calculation techniques: using formulas to calculate mean, variance, and standard deviation, and using a frequency table to find mode, median, and quartiles.

3. Results and Conclusions

The result is a full-fledged assignment for which the student receives automatic feedback. While compiling the STACK assignment and writing the code, the authors encountered some problem areas related to the graphical representation through the `plot()` function.

Issues resolved include:

- A method for constructing bar charts using the `plot()` function through a list of points that are sequentially connected by segments.
- Plotting graphs of piecewise given functions using the `plot()` function through a list of individual functions $y = F(x)$, acting on its segment.
- Line and dot style for function list and point list
- Integrate graphs into multiple choice questions to display images instead of text. Technically, this was implemented by adding images to the answer selection options.
- Random mixing of graph options at each iteration. This measure was introduced to reduce the likelihood of guessing the correct answer. Implementing random shuffling ensures fair testing and encourages students to develop a deeper understanding of the material.
- Determine the correct answer according to the graph PDF and CDF.

Despite this, there are still unresolved issues related to the appearance of the CDF plot due to the point and line style options. For example, the line style display option is not of the dashed line type, which makes it impossible to apply the necessary dashed lines to indicate a break in the CDF plot.

To achieve the current results, the practice of breaking up and simplifying large objects (for example, dividing a single task into parts or dividing a table into rows) was introduced, which made it easier to customize overall feedback and prompts in a more targeted and effective manner.

In the future, the authors intend to explore the possibility of automatically generating or modifying task text and providing personalized feedback using a potential response tree (PRT). The most complex form of feedback is explicit feedback (EF), which often provides an explanation for why a given answer is (in)correct or provides metacognitive cues about how to proceed (Shute, 2008; Van der Kleij et al., 2015).

Bibliography

Aalto OpenLearning. (n.d.). *Randomly shuffling the options*. Retrieved May 30, 2024, from https://openlearning.aalto.fi/question/type/stack/doc/doc.php/Authoring/Multiple_choice_questions.md.

Chongchitnan, S., & Harrison, K. (2021). *Real Analysis and STACK*. International Meeting of the STACK Community 2021. https://doi.org/10.5281/zenodo.5039892

Departament of Mathematics and Statistics. *Plotting*. Retrieved May 30, 2024, from https://maths.cnam.fr/Membres/wilk/MathMax/help/Maxima/maxima_8.html

Forbes, C., Evans, M., Hastings, N., & Peacock, B. (2011). *Statistical distributions*. John Wiley & Sons.

HFT Stuttgart Moodle. (n.d.). *Plots and graphics in STACK questions*. Retrieved May 30, 2024, from https://moodle.hft-stuttgart.de/question/type/stack/doc/doc.php/CAS/Plots.md

Hooper, C., & Jones, I. (2023). *Conceptual Statistical Assessment Using JSXGraph*. International Journal of Emerging Technologies in Learning (iJET), 18(01), pp. 269–278. https://doi.org/10.3991/ijet.v18i01.36529

Härterich, J. (2019). *Using randomized quizzes in undergraduate linear algebra and multivariable calculus*. https://doi.org/10.5281/zenodo.2582874

Knaut, J., Altieri, M., Bach, S., Strobl, I., & Dechant, K. (2022). *A Theory-Based Approach of Feedback in STACK-Based Moodle Quizzes Taking into Account Self-Regulation and Different Proficiency of Learners*. International Journal of Emerging Technologies in Learning (iJET), 17(23), pp. 38–55. https://doi.org/10.3991/ijet.v17i23.36425

Kröger, S., & Schwarz, P. (2020). *Using STACK with Moodle in the Physics Laboratory to review the results of student experiments and their evaluation*. https://doi.org/10.5281/zenodo.4415620

Lutz, T. (2019). *GeoGebra and STACK. Creating tasks with randomized interactive objects with the GeoGebraSTACK_HelperTool*. https://doi.org/10.5281/zenodo.3369598.

Mertens, U., Finn, B., & Lindner, M. A. (2022). *Effects of computer-based feedback on lower- and higher-order learning outcomes: A network meta-analysis*. Journal of Educational Psychology, 114(8), pp. 1743–1772. https://doi.org/10.1037/edu0000764.

Obilor, E.I. (2019) *Feedback and Students' Learning*. International Journal of InnovativeEducation Research 7(2):40-47, April-June., 2019. https://www.researchgate.net/publication/343609551_Feedback_and_Students'_Learning

Oulu University Moodle. (n.d.). *JSXGraph*. Retrieved May 30, 2024, from https://moodle.oulu.fi/question/type/stack/doc/doc.php/Authoring/JSXGraph.md

Riebe, K., & Varmaz, A. (2021). *STACK in Econometrics*. International Meeting of the STACK Community 2021. https://doi.org/10.5281/zenodo.5036033

Safiulina, E., & Labanova, O. (2020). *STACK: stages of the process through the beginner's experience*. https://doi.org/10.5281/zenodo.3955597

Safiulina, E., Labanova, O., & Šeletski, A. (2023). *Geometry in Engineering with STACK automatic assessment*. International Meeting of the STACK Community 2023, Tallinn, Estonia. https://doi.org/10.5281/zenodo.8032713

Safiulina, E., Labanova, O., Seletski, A., & Babenko, K. (2021). *Step-by-step Solution of Applied Engineering Problems in Mathematics Courses Using STACK*. International Meeting of the STACK Community 2021. https://doi.org/10.5281/zenodo.5035975.

Schmitt, K. & Spatz, V. (2020). *Digital Support for physics service lectures*. https://doi.org/10.5281/zenodo.3945813

Shute, V. J. (2008). *Focus on formative feedback*. Review of Educational Research, 78(1), 153–189. https://doi.org/10.3102/0034654307313795.

Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). *Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis*. Review of educational research, 85(4), pp. 475-511.

Wild, C. (2006). *The concept of distribution*. Statistics Education Research Journal, 5(2), 10-26.

Woollett, E. L. (2014). *Maxima by Example: Ch. 2, Plots, Files, Read, Write, and Fit*. Retrieved May 30, 2024, from https://home.csulb.edu/~woollett/mbe2plotfit.pdf

Information on the authors

Elena Safiulina

Elena Safiulina (Ph.D. in Mathematics) is a lecturer in mathematics at the Centre for Sciences, TTK UAS. She has experience with diverse e-learning environments and interactive multimedia learning tools in the field of STEAM education.

Natalja Maksimova

Natalja Maksimova, with master's degrees in mathematics education and business IT, teaches at TTK UAS, and has 16 years' experience in education and e-course development.

Oksana Labanova

Head of Centre for Sciences, TTK, MSc in Mathematics. Experience in ICT tools for teaching. State committee member for e-course evaluation and math assessment. Conducts STEM teaching seminars, focuses research on interactive materials.

Focuses: Implementation and usage of GeoGebra, JSXGraph, or other Programs; New users and authoring of questions.

Article number: 06

GeoGebra in STACK. How it is done and what you need to consider for calculating feedback

Tim Lutz*

University of Kaiserslautern-Landau

Abstract

GeoGebra in STACK makes it easy to integrate GeoGebra elements into STACK tasks. GeoGebra in STACK was designed and developed by Tim Lutz and is integrated into STACK from STACK 4.5 and can be used without additional installation effort. The article is divided into 3 sections: (i) Overview of teaching and learning material to create "GeoGebra in STACK" tasks, (ii) The Kart Race task as an example of the combination of GeoGebra and STACK. (iii) Considerations for the evaluation of GeoGebra in STACK tasks using the example "angle".

* Corresponding author: tim.lutz@rptu.de

1. How to create “GeoGebra in STACK” tasks

GeoGebra in STACK tasks can be used from STACK 4.5 (Lutz, 2023).If you have an older STACK version and cannot change it, you can use the GeoGebraSTACK_HelperTool, which works slightly differently and is not quite as convenient, but is cloud-independent by default (Lutz, 2019).GeoGebra in STACK is designed as a block element, a standard for extending the core functionalities of STACK. The software is firmly integrated into STACK so that no additional installation effort is required. GeoGebra in STACK was designed and developed by Tim Lutz.

Important places to get started with GeoGebra in STACK

On GitHub you can find the documentation of GeoGebra in STACK for developers and for task creators there is a section in the documentation for authoring tasks:
https://github.com/maths/moodle-qtype_stack/blob/dev/doc/en/Authoring/GeoGebra.md

A complete beginner's workshop in GeoGebra in STACK can be found as a video recording from the event “Learning Mathematics: Digital and Interactive”:
https://tim-lutz.de/GGBappletsIntoSTACK/

Basic knowledge of creating STACK tasks is required. (For a more general introduction to STACK (without GeoGebra) see STACK documentation “authoring” section). For a successful understanding and implementation of the example tasks worked on in the workshop, this previous knowledge is not necessary, but only GeoGebra in STACK specific contents are further explained. More information about features of GeoGebra in STACK is available at Lutz (2023).

2. GeoGebra in STACK: New sample task “Kart Race”

New sample tasks were presented at the international STACK Conference 2024 in Amberg. The applications are aimed at getting previously unnoticed user groups interested in working with GeoGebra in STACK.

This contribution is intended to show that the focus can be set not only on students and the use of STACK in higher education, but also on its use for schoolchildren. The graphical representations in GeoGebra offer a wide range of potential applications far beyond solving equations and working with terms.

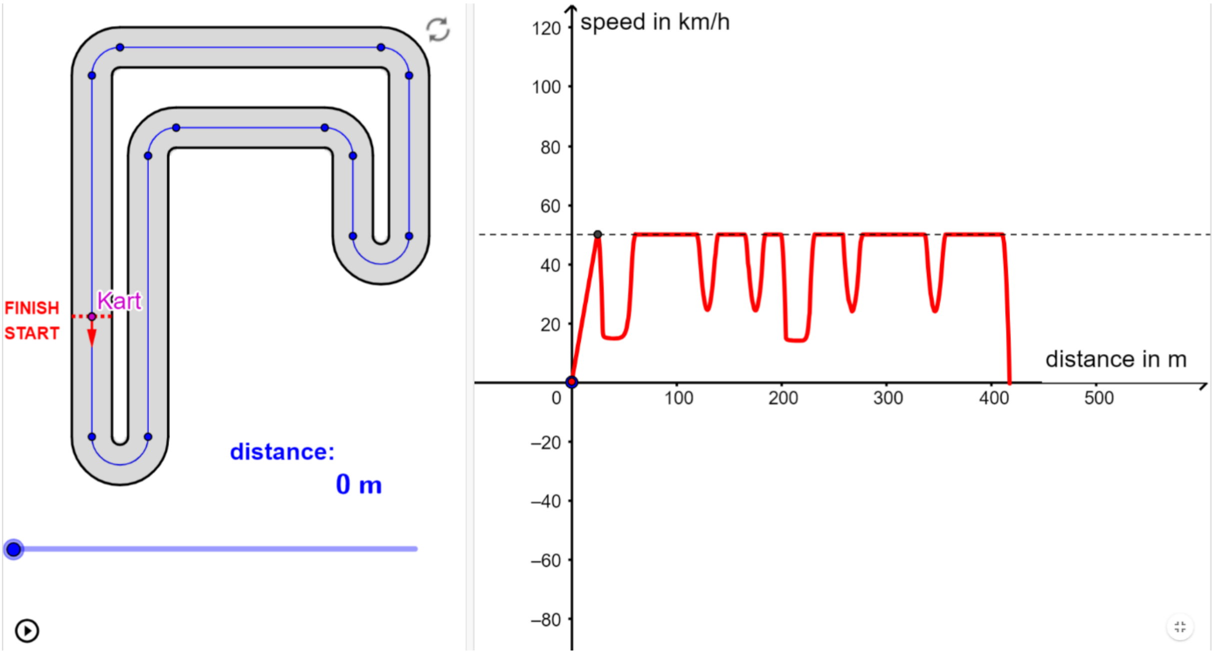


Figure 1: Kart-Race, an applet derived by Tim Lutz from “Kartrennen” by Jürgen Roth

The “Kartrennen” application by Jürgen Roth has been slightly modified to create a STACK question using GeoGebra (https://www.geogebra.org/m/jcvssr2a). In the STACK task developed by Tim Lutz several subtasks must be completed.

Task download link: see workshop materials.

What is the maximum speed of the kart? To determine this value, the graph must be observed in the GeoGebra view (global maximum).

How much distance has it traveled up to the middle of the first bend? To determine the answer to this question, one must either read the graph or, more simply, move the kart into the first curve using the slider. The question of the length of the complete route can also be solved by reading the graph or moving the kart.

Move the kart to the place where it is slowest (not at start/finish). This task can only be completed graphically and interactively. The submission consists of the one-dimensional position of the cart (value “s” in GeoGebra).

How slow is it at the slowest point (not at start/finish)? To answer this question, the graph must be read again.

The “Kart-Race” example tasks shows features that are also suitable for a younger user group. If algebraic tasks are reduced to arithmetic tasks and graphically interactive elements are added using GeoGebra, STACK tasks also become more attractive for use in schools. GeoGebra offers the option of dynamically linking several views. This means that you can always see the race track and graph at the same time, as well as the linked position of the kart in both views.

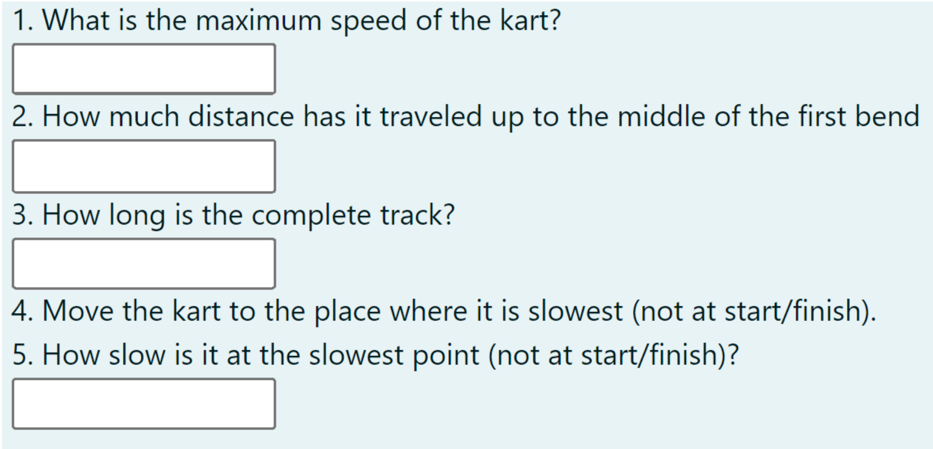


Figure 2: Kart-Race inputs

For many tasks created in this way, in which the diagram is manipulated, a decision must be made on how to distinguish between a correct and an incorrect submission. This decision is made in STACK Feedback via an answer test. As already described for the GeoGebraSTACK_HelperTool (Lutz, 2019), there are two possible approaches. Either one fixes points on the coordinate system in GeoGebra and check the coordinates for algebraic equivalence or one must allow floats in input and has to define an interval in which answers are accepted as correct by selecting the NumAbsolute or NumRelative answer test, depending on the task.

In the case of the “Kart-Race” task, it must be noted that positions should essentially be set correctly regardless of their relative position on the track, so NumAbsolute should be selected, which can be set globally as an input tolerance.

3. “GeoGebra in STACK” specific considerations for designing tasks

Many of the logical considerations described in Lutz (2019) still have to be made, even though the workflow specified in Lutz (2019) no longer has to be as extensive as with the GeoGebraSTACK_HelperTool.

The first thing to consider is where and with what intention a GeoGebra file should be added to STACK. It is particularly important to determine which variables must be passed on to GeoGebra (e.g. for the purpose of graphical randomization) and which variables must be passed back from GeoGebra to STACK, e.g. the position or values of certain objects such as points, sliders or angles. In order to be able to make the latter decision, it must first be clarified how the evaluation and feedback generation is to take place.

The following example assumes knowledge of the “watch” and “remember” features in GeoGebra in STACK.

Example: 3 points A,B,C can be moved freely. It is to be checked whether a certain angle, e.g. ABC between these 3 points, has a certain value of 200 degrees.

Option 1: You specify that the coordinates of the 3 points are passed on to STACK (A,B,C in the watch section). The angle is then calculated in the feedback variables with the help of maxima using appropriate formulas such as cos. Then it is compared with a teacher answer in the feedback trees using suitable value tests.

Option 2: Add a GeoGebra angle object “a” to the 3 points. Consider if “a” should be invisible in GeoGebra and hide “Algebra view” in GeoGebra. Specify that only the value of the angle should be passed on to STACK ("a" in watch section). To restore student processing after submission, add A,B,C only as objects in the "remember" section. In the feedback variables, the angle, which is passed on to STACK as a radian measure by default, should be converted into degrees and should then be compared with the teacher answer in the feedback trees.

The description of both options makes it clear that there are many ways to implement one and the same task as a GeoGebra in STACK task. Option 2 seems more elegant at first glance, as it requires less code. Skills that the task creator has in GeoGebra can help him to implement the same task more easily, because GeoGebra offers many geometric analysis objects (such as angle sizes).

At the same time, however, Option 2 significantly restricts the information available from the graphical task processing in STACK. If Option 2 is followed, it is not easy to determine whether the angle ACB rather than the angle ABC has the required size, i.e. whether just the role of the points has been reversed by students.

There are therefore not only many ways to implement a task, but the choice of procedure for evaluation and feedback generation also limits the adaptive design of the feedback to a greater or lesser extent.

The considerations on structuring the communication between GeoGebra and STACK already described in (Lutz, 2019) also apply to GeoGebra in STACK:

Which information must be communicated to GeoGebra by STACK in order to display the task? (set)

Which information should GeoGebra pass on to STACK in order to calculate feedback in STACK? (watch)

Which information should be restored in GeoGebra after changes have been made by the user after the task has been submitted? (watch/remember)

Bibliography

Lutz, T. (2019). *GeoGebra and STACK. Creating tasks with randomized interactive objects with the GeoGebraSTACK_HelperTool*. https://doi.org/10.5281/zenodo.3369599

Lutz, T., Pinkernell, G., & Sangwin, C. (2023). *GeoGebra integrated in STACK*. International Meeting of the STACK Community 2023, Tallinn, Estonia. https://doi.org/10.5281/zenodo.8133499

Information on the author

Tim Lutz

Tim Lutz is a PostDoc at the University of Kaiserslautern-Landau. As an employee in an Erasmus project of the PH Heidelberg and University of Edinburgh, he designed and developed the software "GeoGebra in STACK" and integrated it into STACK. Tim Lutz regularly holds workshops for various user groups.

Focuses: STACK in teaching or exams; Other topics related to STACK.

Article number: 07

Implementation and effectiveness of mathematical STACK problems in higher education for engineers: Lessons learned.

Oleg Boruch Ioffe*, Maike Schelhorn, Jessica Schäfer, Reik V. Donner

h² – Hochschule Magdeburg-Stendal; Germany

Abstract

For several years, Magdeburg-Stendal University of Applied Sciences has been using digital mathematics problems for enriching its mathematics teaching. With the different study areas and lecturers, the teaching formats and the respective degree of STACK integration differ markedly among the courses that already use STACK problems. This broad range of settings allow drawing a comprehensive picture of different strategies for implementing STACK into higher education programs for future engineers.

By means of systematic student questionnaires and individual qualitative interviews organized by social scientists, we obtain detailed data on potential factors that may affect the students' acceptance of STACK problems as a precious resource of information on utilization of learning materials as well as individual learning success. In addition, we perform detailed learning analytics based on access data from the learning management system "moodle" along with success rates of the final course exams. Combining all those complementary data sources, we are in a unique position for determining factors controlling the acceptance and effectiveness of STACK problems in higher mathematics education for engineers.

* Corresponding author: oleg-boruch.ioffe@h2.de

1. Introduction

Since the beginning of the summer term of 2022, we have been using digital problems as part of the course "Mathematics 2" in the Civil Engineering program of Magdeburg-Stendal University of Applied Sciences. These problems have been initially developed using the commercial software WIRIS during the Covid-19 pandemic and have been gradually replaced by STACK over the last few years. All STACK problems are collected within a resolute moodle course "Mathematics Learning Support Centre (MLSC)", where they are being checked for their didactic and technical implementation properties before practical usage. They are then available to all interested university students in various categories, such as fundamentals of higher mathematics, linear algebra, and calculus. More information on the corresponding details can be found in Donner (2023).

2. Use of digital STACK problems in university teaching of engineering mathematics

At the international STACK conference held in Tallinn in 2023, we already presented our initial findings and experiences regarding the advantages and disadvantages of working with the two different systems WIRIS and STACK (Judakova 2023). Although WIRIS provides a useful graphical interface, the decisive factor for us leading to the decision to switch to STACK was that WIRIS does not have a large community and requires a paid license. By contrast, we take strong benefits from the large international STACK community and the fact that STACK is open source.

Most of the STACK problems that we currently use at our university have been originally taken from the DOMAIN database, which is coordinated by the Ruhr University of Bochum. In addition, a second digital problem collection of the Technical University of Cologne has been partly exploited as well. Many of these STACK problems had to be checked by our team for technical functionality first and have only been utilized as study material after any errors had been corrected. Furthermore, syntactic and didactic adjustments in relation to the notation of mathematical expressions used at the university had to be applied to the original STACK problems before they were incorporated into the respective courses. Since not all areas of the curriculum were equally covered by the initial collection, we additionally exploited some STACK problems from the HELM workbooks and translated them into German accordingly. If individual topics are still not available in any of the aforementioned databases (such as applications of l'Hospital 's rule), new STACK problems are being developed and implemented in STACK.

3. Use of STACK problems in mathematics courses

The use of STACK problems within regular courses took place for the first time in the winter term of 2023/2024, where we employed the problems as individual training materials and for voluntary e-assessments in the course “Mathematics 1" and partly "Mathematics 3" in the Civil Engineering program. In the following summer term of 2023, STACK problems were first exploited in the courses "Mathematics 2” for Civil Engineering and "Mathematics 1” in the bachelor program “Sustainable Resources Economics and Management" (StREaM). In the winter term of 2023/2024, the STACK problems were fully integrated in the courses "Mathematics 1-3" (Civil Engineering), "Mathematics 2” (StREaM) and "Mathematics 1” (Water Engineering and Recycling Management).

From this successive increase in the number of courses supplemented by STACK problems, it can be deduced that our initial successful exploitation of STACK in the context of the mathematics courses in the bachelor's degree program on Civil Engineering convinced other mathematics professors and lecturers to use STACK in their mathematics courses as well. Meanwhile, STACK problems have become an integral part of the courses Mathematics 1-3 for Civil Engineering (in German), Mathematics 1-2 for StREaM (English) and, starting in fall 2023, also Mathematics 1-2 for Water Engineering and Recycling Management (German), which are currently taught by three different lecturers from the university's two engineering faculties. Full coverage of all introductory (first year) mathematics courses in all engineering programs is planned for fall 2024, when STACK problems shall also be rolled out as part of the mechanical and electrical engineering as well as engineering economics programs.

4. Exemplary teaching design: Mathematics 2 for Civil Engineering

In the following, we will report our experiences with the STACK implementation in the course "Mathematics 2" for Civil Engineering as an example. Of 14 to 15 term weeks, 13 weeks are usually reserved for lectures and practical sessions supported by self-study periods for intensive individual work with the teaching materials. In the specific case of Mathematics 2, the students experience the lectures in an inverted classroom model and are expected to study video recordings of teaching sequences before attending the respective lectures and practical sessions. Weekly electronic self-assessments utilizing STACK problems are topically aligned with all other lecture materials. At intervals of around 3-4 weeks, additional digital moodle quizzes based on selected STACK problems are used as voluntary e-assessments that conclude the respective topic complex and allow students to earn bonus points for the final exam. For example, in the course "Mathematics 2", the following four bonus quizzes are offered:

- Bonus quiz 1: Sequences and series, general properties of real-valued one-variable functions.
- Bonus test 2: Classes of real-valued functions and solving of nonlinear equations.
- Bonus test 3: Differential calculus and its applications.
- Bonus test 4: Integral calculus and its applications.

In the other courses enriched by STACK problems, the teaching formats vary depending on the study area and responsible lecturer, including courses with face-to-face lectures and recitations, face-to-face lectures with integrated examples, and lectures in the inverted classroom combined with face-to-face recitations. There are also clear differences between these courses in the respective degree of STACK integration, such as curated collections of tasks for self-study, weekly digital exercises and/or integrated e-assessments.

5. Systematic student questionnaires and first results of data collection

Since the beginning of the 2022/23 winter term, our efforts for a successive STACK integration have been accompanied by regular evaluations by social scientists. Commonly, three surveys are conducted per term among the participating students in the form of structured questionnaires, where the originally paper-based questionnaires have meanwhile been replaced by digital surveys. At the beginning of each term, students are initially asked about their own assessments of their expected knowledge and mathematics as well as general examination related fears. Mid-term and final surveys provide further insights into the development and change in the self-assessment of the respective study group during the course of as well as at the end of the term. In addition, the evaluation staff conducts qualitative interviews with selected students at the end of the term.

For interpreting the results of STACK based e-assessments, written exams as well as student surveys in the Civil Engineering problem, we have to distinguish between two parallel study programs – a regular one with classical academic teaching in each term and a dual program where students spend each summer term in a company for practical training. Consequently, the Mathematics 2 course is taught in each semester, which regular (dual) program students participating in each summer (winter) term. To allow for a unique assignment of individual students to any of the two programs, all surveys include one question on the current number of terms in the program and its type (regular or dual).

The first results from our surveys demonstrate that most students, around at least 90%, in both programs highly appreciate the digital STACK problems with direct feedback on the provided solutions. The positive specific feedback on the bonus quizzes, i.e. interim e-assessments allowing to acquire bonus points for the final exam, is also exceedingly high, averaging over 80%. According to the evaluation of the final surveys, the format of voluntary bonus tests is more strongly accepted than the voluntary digital weekly training quizzes, with 83% as compared to 74%.

6. Summary

Initial results of learning analytics based on the digital learning management platform moodle and structured student surveys confirm that STACK problems offered especially in the form of bonus quizzes, but also for training purposes are highly appreciated by the students and contribute positively to their overall learning success. Despite initial efforts for integrating STACK into the different course programs being considerably high, we are confident that this effort is justified by the gained results. Additional lecturers have already joined our initiative for integrating STACK into the mathematics courses of other engineering programs; a complete roll-out in the all-study programs of the two engineering faculties of Magdeburg-Stendal University of Applied Sciences is expected to become effective in the 2024/25 winter term. Corresponding quantitative results of systematic assessments on the learning performances and students' self-assessments across all corresponding programs will be reported in future publications.

Acknowledgments

The presented study was funded by the “Stiftung Innovation in der Hochschullehre” as part of the project “h²d²: didaktisch und digital kompetent lehren und lernen“.

Bibliography

Donner, R.V.; Judakova, G.; Ioffe, O.B.; Brandt, K.; König, L. (2023). *Das digitale Mathematik-Lernzentrum der Hochschule Magdeburg-Stendal und seine Integration in die Grundlagen-Lehrveranstaltungen Mathematik* (in German). In Liebscher, E. et al. (eds.): Digitale Lehre im Rahmen der Grundlagenausbildung in MINT-Fächern an Hochschulen. Tagungsband, Merseburg.

Ioffe, O. B., Judakova, G., Schäfer, J., Hajji, R., König, L., Donner, R.V. (2023). *Usage scenarios of mathematical STACK problems at the Magdeburg-Stendal University of Applied Sciences*. Contributions to the International Meeting of the STACK Community 2023, 24.-

26.04.2023, TTK University of Applied Sciences Tallinn, Estonia. DOI: 10.5281/zenodo.8032250

Judakova, G., Ioffe, O. B., König, L., Donner, R.V. (2023). *A Digital Mathematics Leaning Support Centre Based on a Curated German-Language Collection of Mathematical STACK Problems*. In: Contributions to the International Meeting of the STACK Community 2023, 24.-26.04.2023, TTK University of Applied Sciences Tallinn, Estonia, DOI: 10.5281/zenodo.8032271

Information on the authors

Oleg Boruch Ioffe

Research Associate at the Mathematics-Statistics Teaching-Learning Laboratory as part of the h²d² project at Magdeburg-Stendal University of Applied Sciences and Lecturer in Mathematics since 2021.

Maike Schelhorn

Research Associate at the Mathematics-Statistics Teaching-Learning Laboratory as part of the h²d² project at Magdeburg-Stendal University of Applied Sciences since 2023.

Jessica Schäfer

Research Associate at the Evaluation Laboratory as part of the h²d² project at Magdeburg-Stendal University of Applied Sciences since 2022.

Reik V. Donner

Professor of Mathematics (Data Science & Stochastic Modelling) since 2018 and head of the Mathematics-Statistics Teaching-Learning Laboratory as part of the h²d² project at Magdeburg-Stendal University of Applied Sciences since 2021.

Focuses: Automation of question and feedback generation; STACK in teaching or exams.

Article number: 08

Investigating the feasibility of Automating the Advanced Higher Mathematics and Physics Scottish school exams

Konstantina Zerva*, Ben Chooyin, Jack Fortune, Rachel Hollingsworth, Mitchell McLachlan, Heather Napthine

School of Mathematics, The University of Edinburgh, Edinburgh, UK

Abstract

Automatic online assessment is becoming increasingly common for formative work in mathematics and other STEM disciplines. Many courses at university level have weekly auto graded quizzes which are used to test students' understanding and give instant feedback. Also, the COVID-19 pandemic accelerated the use of online assessment in high stake summative exams. There are several advantages to automating examinations, including increased efficiency, reduced grading bias, and improved feedback for students.

The aim of this paper is to investigate the feasibility of automating the SQA (Scottish Qualification Authority) Advanced Higher Mathematics and Physics examinations using STACK. A past exam paper for each course was implemented into STACK. While the majority of the questions were successfully automated in STACK, there are questions for which the direct automation is not possible. However, upon closer examination of the more challenging questions, it became obvious that many of them had underlying objectives that could be effectively assessed using STACK.

Subsequently, we successfully directly translated 81% of the marks from the SQA's 2019 Advanced Higher Mathematics paper and 69% of the marks from the SQA's 2019 Advanced Higher Physics paper into an automated format. These achievements highlight the promising potential for further research in this area which may allow for the full automation of these examinations.

* Corresponding author: K.Zerva@ed.ac.uk ORCID: 0000-0002-1914-6187.

1. Introduction

Over the past 30 years, advancements in Computer Aided Assessment (CAA) of Mathematics and other STEM disciplines have significantly transformed the educational landscape. Until recently, automatic assessment systems were reliant on multiple choice questions (MCQ) or similar response question types. Contemporary CAA systems accept answers which contain mathematical content and establish the mathematical properties of those answers using computer algebra. The systems not only provide a score, but can also provide immediate feedback on the student's attempt.

At the university level, the integration of automated assessments is becoming increasingly essential. At the University of Edinburgh, we systematically use CAA with STACK for the majority of first- and second-year mathematics modules and for many other mathematics and general science courses, since 2016. These assessments are mainly formative, and our students have benefited by having (i) immediate feedback; (ii) consistent criteria and (iii) repeated practise of randomly generated questions (C. Sangwin & Zerva, 2020). We have been gradually using STACK in more high-stakes situations such as the fully online courses "Fundamentals of Algebra and Calculus" (FAC) and "Introductory Mathematics with Applications" (IMA) (Gratwick et al., 2020; Kinnear et al., 2021; Kinnear, 2018), and for running a fully online exam for our first-year course "Introduction to Linear Algebra" (ILA) (C. Sangwin, 2019).

While STACK is mainly used as an assessment tool in Higher Education, it was also found that it would be suitable for automating questions from existing school level exams, more specifically for the International Baccalaureate examination (C. Sangwin & Köcher, 2016).

The work reported in this paper is based on an undergraduate mathematics project which the first author supervised during the academic year 2023 – 2024. The Mathematics Project is a final year 20 credit course at level 10 according to the Scottish Credit and Qualifications Framework (SCQF), where students undertake a group project to investigate a particular mathematical subject in depth. The project was inspired by the work reported by (C. Sangwin & Köcher, 2016). The aim of the project is to explore the feasibility of utilising STACK for automating the Scottish Qualification Authority (SQA) Advanced Higher Mathematics and Physics examinations.

This paper is organised as follows. Section 2 provides some background information. Section 3 defines our methodology for evaluating the extent to which questions can be automatically marked. Results are given in Section 4, with a discussion following in Section 5. Section 6 summarises the main conclusions.

2. Background

The Scottish Qualifications Authority

Established in 1997, the Scottish Qualifications Authority (SQA), https://www.sqa.org.uk, serves as the national body responsible for overseeing the majority of qualifications in Scotland, excluding university degrees. The SQA offers a wide range of subjects, including Mathematics and Physics at the Advanced Higher level. In 2022, the Advanced Higher Mathematics examination was taken by 3,915 candidates - while 2,130 students sat the Advanced Higher Physics examination (Stephen, 2022).

The Scottish Qualifications and Credit Framework (SCQF), https://scqf.org.uk, is a part of the SQA that provides a framework for understanding the difficulty level of a given qualification.

Advanced Higher courses are classified at SCQF Level 7, while the first year of university typically encompasses SCQF Levels 7 and 8. The Advanced Highers Scottish national exams are equivalent to the English A Levels exams and to International Baccalaureate.

The Push for Digital Examinations

In an era increasingly dominated by digital technology, with devices like iPads and computers becoming central to both educational and professional settings, the push towards digitalisation in education reflects a broader societal shift. Embracing digital examinations could facilitate the integration of these technologies into the curriculum. This approach not only aligns with the technological realities of the modern world but also offers a learning environment that reflects the environments that students are likely to encounter in their future careers.

Scotland's Curriculum for Excellence (CfE), https://education.gov.scot/curriculum-for-excellence is a national educational framework designed to provide a cohesive and flexible approach to learning for children and young people aged 3 to 18. CfE promotes the use of technology in classroom even for children at nurseries (ages 3-5). This includes familiarising young children with basic digital skills, promoting safe and responsible use of technology, and using digital resources to support creativity, problem-solving, and collaboration. All pupils from Primary 6 (age 9-10) to Secondary 6 (age 16-18) receive their own digital device (usually an iPad) from school, to have personal access to digital learning with their teacher in school or at home.

As we move further into the digital age, it appears inevitable that digital examinations could become the standard. Whilst Scotland has not announced any plans for digital examinations, proposals by examination boards in England suggest that at least one major subject may be offered in a partly digital format by 2030 (Thake, 2023). Should England move forward with digital examinations, Scotland may follow suit. The shift towards digital examination requires careful consideration of the benefits and drawbacks, as well as significant investment in advanced software to ensure a smooth transition to this new format.

Digital examinations offer a range of practical advantages that could enhance the educational experience. Unlike the traditional method, where students endure a lengthy wait for their grades, digital examinations can deliver results almost immediately. Digitalisation would make the creation and grading of examinations more efficient and could reduce subjective biases that affect grading, such as those related to handwriting. At a macro level, digital examinations offer government bodies the ability to easier conduct large-scale analysis of examination data. Moreover, the transition to digital format would be environmentally friendly; the traditional paper-based assessments have a significant ecological footprint, with millions of sheets of paper consumed annually for SQA examinations ("Disclosure Log", 2023).

This shift is not without its challenges. The risk that digital examinations may advantage students with ready access to technology at home cannot be overlooked. Digitalisation could exacerbate existing educational inequalities, particularly affecting students from lower-income families. Also, digitalisation may raise concerns regarding security and data protection. Ensuring robust cybersecurity measures are in place is of utmost importance in protecting sensitive information from unauthorised access. Furthermore, the reliance on technology introduces the potential for technical difficulties which could disrupt the examination process (e.g., poor broadband in isolated areas).

With this paper we do not want to advise the examination boards about the way they conduct the exams. We only want to investigate if, with the existing technology, it is possible to recreate a digital version of the examinations that accurately distributes the designated marks.

3. Methodology

There are many CAA systems which could have been chosen for this study and would have given similar outcomes. We used the STACK software (System for Teaching and Assessment using a Computer algebra Kernel) https://stack-assessment.org/ which is the main assessment system we use at the school of Mathematics. STACK is designed to enable students to answer questions by typing mathematical expressions and the answers are graded based on the mathematical properties of these expressions. STACK forms a question type in the Moodle Learning Management System. We only used STACK for this work, but depending on the nature of the questions other Moodle question types could be used such as Pattern Match and Essay.

Automating the Advanced Higher Mathematics Examination

Our exploration began by examining the available Advanced Higher Mathematics past papers accessible via the official SQA website https://www.sqa.org.uk/pastpapers/findpastpaper.htm.

These papers span from 2018 onwards, with the exception of the years 2020 and 2021 due to the disruptive effects of the COVID-19 pandemic. Our analysis revealed remarkable consistency across these papers, with the questions largely retaining the same structure and format from one year to the next. Given the consistent nature of the past papers, we concluded that successfully automating a single paper would suffice to demonstrate the viability of using STACK for Advanced Higher Mathematics examinations. Consequently, the 2019 paper ("SQA 2019 Advanced Higher Mathematics Past Paper", 2019) was selected. The paper has 17 questions which contribute to 100 marks in total.

The official marking instructions are also available in the SQA website ("SQA 2019 Advanced Higher Mathematics Finalised Marking Instructions", 2019). Reading the marking instructions, it is clear that the students' working is of equal importance as to the accuracy of the final answer. It is stated that: "Only award full marks where the solution contains appropriate working. A correct answer with no working receives no mark, unless specifically mentioned in the marking instructions."

For each question we undertook the following evaluation:

1. Can the question be automatically marked completely with STACK? All marks must be assigned exactly as in the mark scheme.
2. Can we assess the final answer(s) automatically and completely with STACK?
3. Can the final answer imply the methods used?

We also paid attention on the form of the answer and the syntax required for entering it into STACK. Since the target audience is school pupils, we wanted to make sure that it would be straightforward for them to type in their answer without requiring to know any specific Maxima syntax.

Automating the Advanced Higher Physics Examination

As with the Mathematics paper, we began by examining the complete range of Advanced Higher Physics past papers accessible via the SQA website. These papers range from 2018 onwards, excluding the years 2020 and 2021. Through our analysis, we noted a striking consistency among the papers, surpassing even that observed in the Mathematics examination. Notably, patterns of similarity and repetition were apparent within each paper as well. Given the uniformity observed across all papers and questions within them, we deemed it sufficient to automate a singular past paper. Accordingly, we selected the 2019 paper to maintain consistency with the Mathematics automation. The paper has 16 questions which contribute to 140 marks in total. For each question we followed similar evaluation as with the Mathematics paper.

4. Results

Automating the Advanced Higher Mathematics Exam

Fully automated questions

There were numerous instances where questions could be automated without any alterations, while still awarding all relevant partial credit for the various mistakes outlined in the official marking scheme (“SQA 2019 Advanced Higher Mathematics Finalised Marking Instructions”, 2019). One such example is Question 2 (b) of the examination paper (“SQA 2019 Advanced Higher Mathematics Past Paper”, 2019), where students are tasked with computing the product of two matrices. In this case, one mark is awarded for each correctly completed row, a setup we successfully implemented. The implementation of this question is depicted on the left of Figure 1 while the official marking scheme on the right. However, we note that there was one difference made to the original question: students are aware that their answer must be a 3×2 matrix since the dimensions are already specified. There is an option to have an answer box of variable size, thereby requiring students to select the dimensions of their answer. However, this has a much less intuitive interface and would require students to do some coding to input their answer. Overall, this adaptation would not be beneficial, as it would unnecessarily complicate the process of entering answers for the student. This is a minor consideration, and the question functions very well as it is.

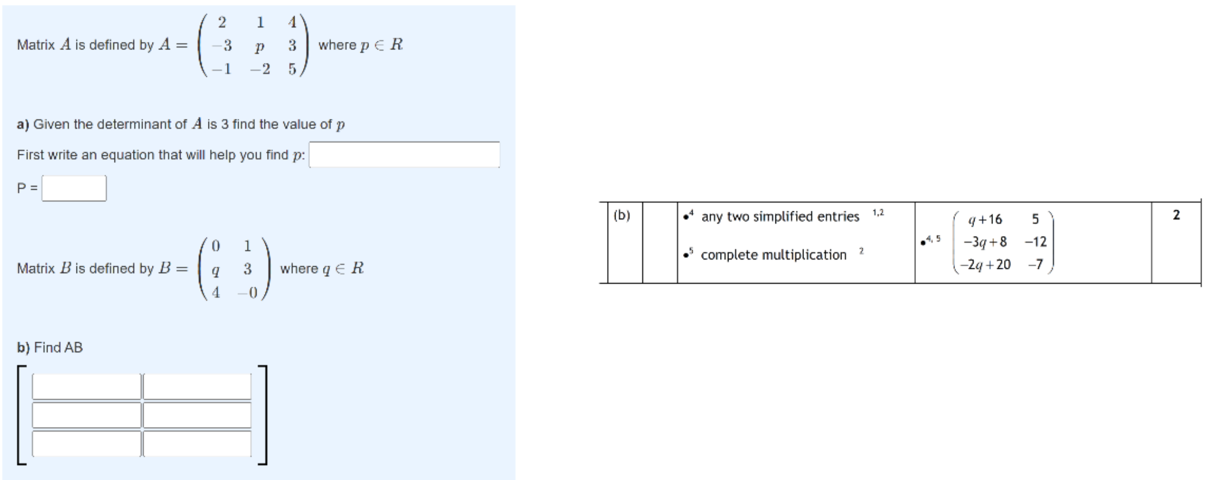
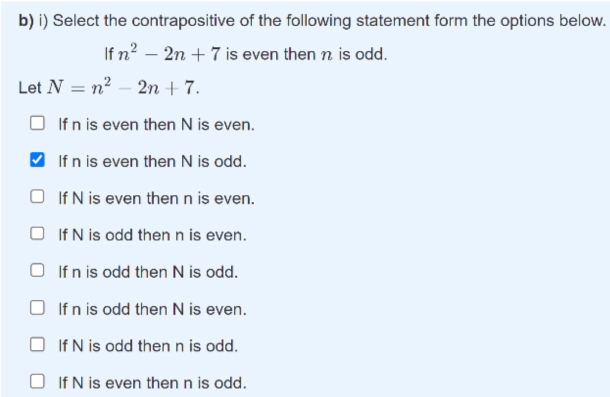


Figure 1: Matrix Question Implementation (left) and SQA official marking scheme (right). The marking scheme is subject to Copyright © Scottish Qualifications Authority.

Several other questions were successfully automated by making minor adjustments to them. Consequently, 56% of the marks from the SQA’s 2019 Advanced Higher Mathematics paper were directly translated into an automated format.

Questions required modifications

Certain questions require modification to be compatible with the STACK system. An illustration of this need for adjustment is found in question 11(b)(i) of the paper. Originally, the question required students to state the contrapositive of a given statement, without providing any possible answers. The answer will involve a mixture of text and formulas which is difficult to type into STACK or assess it properly. We decided to introduce multiple-choice options that cover all possible formulations of the correct contrapositive. This change does not oversimplify the question and still demands that students recall and apply the procedure for inverting a statement to its contrapositive. However, by providing choices that define the concept, students are able to earn full credit where they may have received none otherwise. Considering the low stakes of the question (worth only 1 mark) and the low probability of a student randomly selecting the correct answer (1/8), we are confident that the adaptation to a multiple-choice format maintains the question's original objectives and integrity without compromise. Figure 2 compares the STACK format for this question and the corresponding marking scheme.


b) i) Select the contrapositive of the following statement form the options below.

If $n^2 - 2n + 7$ is even then n is odd.

Let $N = n^2 - 2n + 7$.

- ☐ If n is even then N is even.
- ☑ If n is even then N is odd.
- ☐ If N is even then n is even.
- ☐ If N is odd then n is even.
- ☐ If n is odd then N is odd.
- ☐ If n is odd then N is even.
- ☐ If N is odd then n is odd.
- ☐ If N is even then n is odd.

(b)	(i)	•[2] write down contrapositive statement [1,2,8]	•[2] If n is even then $n^2 - 2n + 7$ is odd	1

Figure 2: Contrapositive Question Implementation (left) and SQA official marking scheme (right). The marking scheme is subject to Copyright © Scottish Qualifications Authority.

A further example of this need for adjustment is found in question 3(b), which instructs the student to sketch a graph. We will discuss more about assessing graphs in STACK in the discussion section.

With these adaptations a further 19% of the marks were successfully converted into an automated format allowing us attribute 75% of the marks.

Upon analysing the remaining questions, it became evident that many required short sentence-based answers. Initially, the use of STACK's inbuilt Levenshtein distance function was considered, but its limitations with response variability made it an unsatisfactory option. However, exploring the automated assessment of short written answers revealed potential solutions, as it will be outlined in the Discussion section. Incorporating these advancements into STACK we automated an additional 6% of the mathematics paper, raising the automation potential to 81%.

Of the remaining 19% of mathematics questions, 9% consisted of 'Proof' questions, which posed unique challenges and for which optimal automation processes have not yet been identified. As will be discussed in the Discussion section the Parsons proof feature offered a partial solution, however, it did not fully match the difficulty level of traditional paper-based questions. This suggests that manual grading might remain necessary for these questions.

The remaining 10% of questions proved too challenging to automate. These questions presented complexities in awarding credit for intermediate steps, and in designing an assessment method that wouldn't inadvertently provide hints through the number of input boxes. Nevertheless, as STACK evolves, automating these types of questions could become feasible.

Automating the Advanced Higher Physics Exam

A systematic review of the Advanced Higher Physics papers allowed us to identify that all questions on the papers could be classified into five distinct categories, as outlined in Table 1. In our automation process, unlike the Mathematics paper, we did not find that any question type could be seamlessly translated into STACK, whilst still preserving exactly the original format of the questions from the paper-based examination.

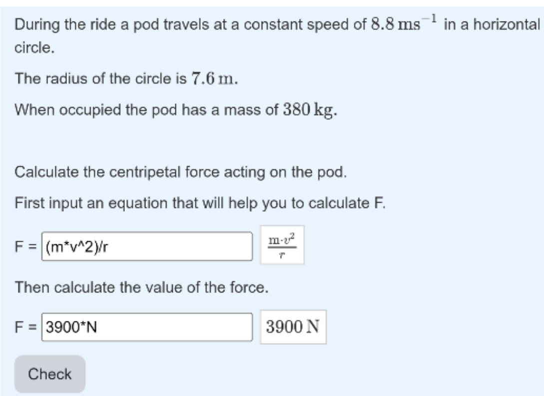
Table 1: Categorising Advanced Higher Physics Questions.

Categorising Physics Questions	
Category	**Description**
Formula and Substitution	Questions requiring candidates to select and evaluate the correct equation.
Short Sentence	Questions requiring a written answer no longer than a couple of sentences.
Diagram	Questions involving drawing and/or interpreting graphs and/or diagrams.
Using your knowledge of Physics	Open-ended questions typically answered with short essay-style responses.
Derivation Questions	Questions requiring candidates to derive known equations.

Formula and Substitution

This type of question was the most common question in the Physics paper. Students were given a problem, and in order to solve it they required to identify the correct formula, substitute the correct values, do the relevant calculations and find the correct numerical answer with appropriate units. The marking instructions for this type of questions gives 1 mark for the correct formula, 1 mark for substituting the values and making the calculations, and 1 mark for the correct final answer with units (see Figure 3, right).

Our proposed method to assess this type of questions was to ask students to type in the correct formula and the final answer (see Figure 3, left), which together constitute around two-thirds of the marks for this question type. These are straightforward to assess using STACK and we coverid 34% of the total marks.



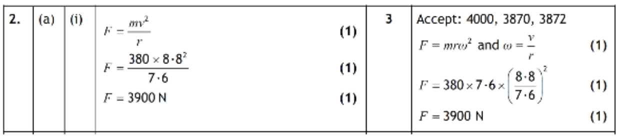


Figure 3: Formula and substitution Question Implementation (left) and SQA official marking scheme (right). The marking scheme is subject to Copyright © Scottish Qualifications Authority.

We also tried to implement this question by using the “Reasoning by Equivalence” feature. As you can see in Figure 4, we tried this without including units (left) and with units (right). While the equivalence is correctly established when units are not included, it fails to establish when we add units. Another drawback of this approach is the specific syntax that is required for substituting the numbers in the equation. We believe that this adds an extra layer of difficulty in the question because it requires pupils to be familiar with the Maxima syntax. For this reason, we decided to keep the previous method for these questions. We acknowledge that our proposed approach still falls slightly short of the SQA’s original objectives of such questions, leaving the optimal method for the assessment of these questions in STACK still undetermined.

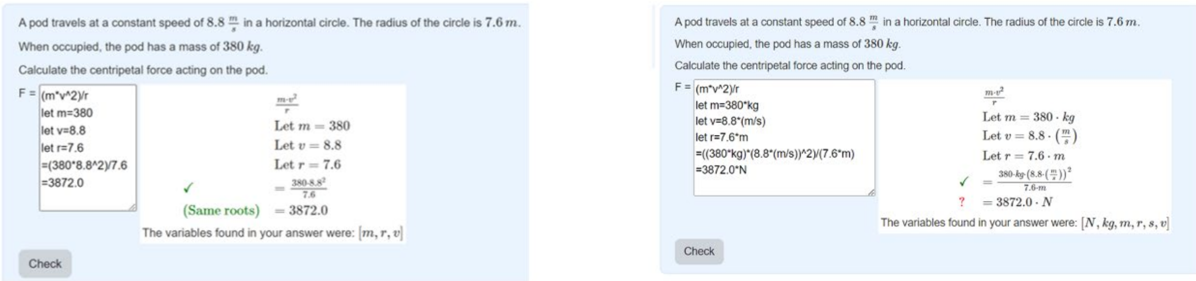


Figure 4: Reasoning by equivalence without using units (left). Reasoning by equivalence with units (right).

Short answer

A large portion of the paper comprised ‘Short Sentence’ questions, and using a similar approach to Mathematics paper we managed to attribute a further 21% of marks. We will discuss more about short answers questions in the Discussion section.

Free body diagrams

Originally, we wanted to use Meclib https://github.com/mkraska/meclib to assess these types of question, something that is suggested in (Orthaber et al., 2020). But it was not possible during the time we run the project to learn how to use the Meclib library alongside learning how to use STACK. So, for this type of question we decided to transform the sketch-based questions into multiple-choice formats. These modified questions should still effectively assess the intended learning outcomes. An illustrative example of this adaptation is the first part of question 2(b) from the 2019 Advanced Higher Physics paper (“SQA 2019 Advanced Higher Physics Past Paper”, 2019), which is presented in Figure 5. By adjusting the free body diagrams to multiple choice questions, we managed to further allocate 14% of the total marks. We acknowledge that this is not an ideal adjustment and we firmly believe that the use of Meclib would have given a more realistic approach to these questions.

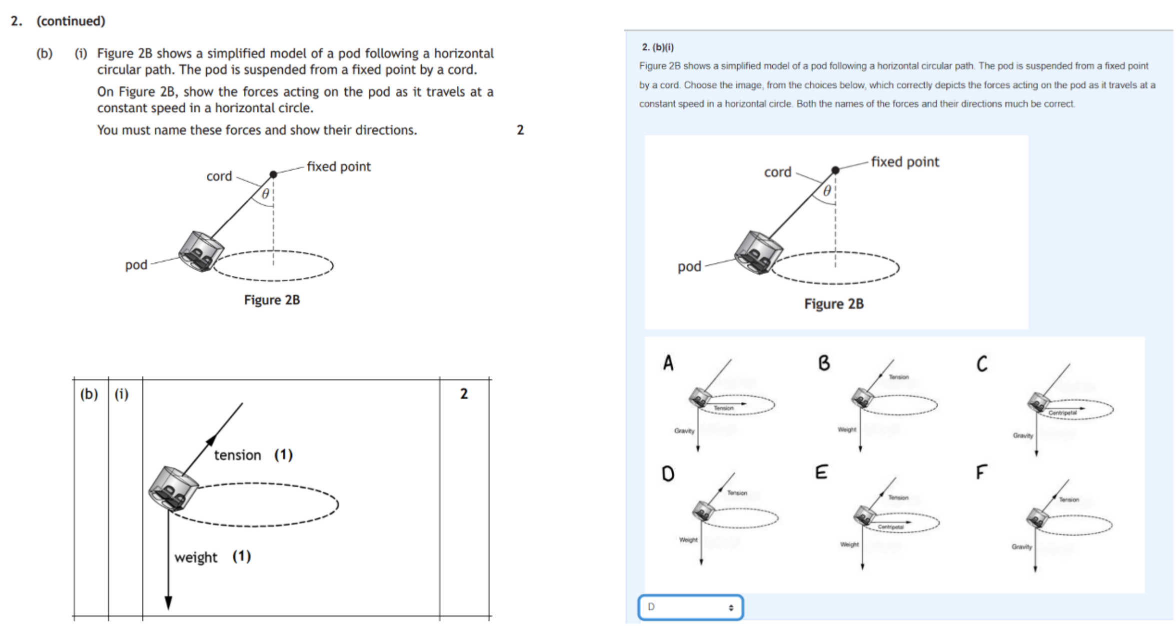


Figure 5: Free body diagram question. The original exam question (left, Copyright © Scottish Qualifications Authority) and the STACK adaptation (right).

Using your knowledge of Physics

With regards to ‘Using your Knowledge of Physics’ questions, we faced significant obstacles. This question type prompts students to provide their understanding of a particular topic in Physics. Students typically approach this question by providing as much theory as possible from any area of Physics that they deem relevant. In order to attain full marks, the student's answer must demonstrate a thorough grasp of correct and pertinent Physics theory.

These questions often blend mathematical formulae with written responses, demanding different evaluation methods. Requesting separate inputs for equations and text could confuse students, especially since equations may be integrated into sentences, and full marks can sometimes be achieved without equations. The absence of a rigid marking scheme and the open-ended nature of these questions further complicate automated assessment.

So, we suggest to manually mark this question type. Despite being somewhat inconvenient, this component accounts for approximately 6 out of the 140 marks on average in a given paper. Therefore, manually grading this minor portion still enables the realisation of the benefits of automation. In this case, the student would complete the entire examination on STACK, where their answer to this question would be marked later by hand.

Derivation

Regarding ‘Derivation’ questions, which have similar format to ‘Proofs’, automatic assessment is particularly challenging due to their inherently complex nature. These questions don’t specify the number of steps required, and providing a set number of response boxes could inadvertently offer hints or restrict students who may have used additional steps. Moreover, some derivations require the use of multiple formulae in any order, further complicating the assessment process. Overall, the unpredictable approaches students might take to complete a given ‘Derivation’ question pose significant challenges to automation.

Thus, the last two categories, the open-ended ‘Using your Knowledge of Physics’ and ‘Derivation’ questions, contribute a respective 4% and 2% to the total marks, and for which effective automated assessment strategies have also yet to be established.

5. Discussion

Our research has revealed that it is highly feasible to fully automate the assessment of Advanced Higher Mathematics. While there are still challenges to overcome regarding the SQA Advanced Higher Physics course, automating its assessment remains an achievable prospect. In this section, we discuss in further depth the primary obstacles encountered during our research.

Graphs

Currently, STACK offers various graph sketching features, yet none exactly replicate traditional paper-based graph sketching. Question 3(b) in the Mathematics paper ("SQA 2019 Advanced Higher Mathematics Past Paper", 2019) instructs the student to 'Sketch the graph of $y = |f(x)|$' for the function $f(x) = x^2 - a^2$. A suggested modification involves transforming the question into a multiple-choice format, as illustrated in Figure 6. To create the various options of the MCQ we took into consideration the various properties of the specific function. The student to recall the definition of the modulus function, denoted by |f(x)|. This function is defined as:

$$|f(x)| = \begin{cases} f(x) \text{ if } f(x) \geq 0 \\ -f(x) \text{ if } f(x) < 0. \end{cases}$$

For values of $f(x)$ below the x-axis, the student must negate them to get the corresponding values for $|f(x)|$ (Fact 1). The student is also required to recall that the (local) maximum turning point must be situated on the y-axis (Fact 2). Additionally, the graph must exhibit line symmetry, and it should not display smoothness at x intercepts (Fact 3). To adapt the question for an online format while preserving its assessment aims, we only need to retain the requirement that the question assesses the student's ability to recall Facts (1), (2), and (3) as defined above.

We acknowledge that this method has inherent limitations. Providing students with multiple choices effectively eliminates the scenario where a student has 'no idea' about the answer. With options available, students will always have a non-zero chance of guessing correctly, although this probability diminishes as the number of options increases, making correct guesses less likely. Additionally, options can unintentionally guide students toward the correct answer by allowing for comparison and informed decision-making.

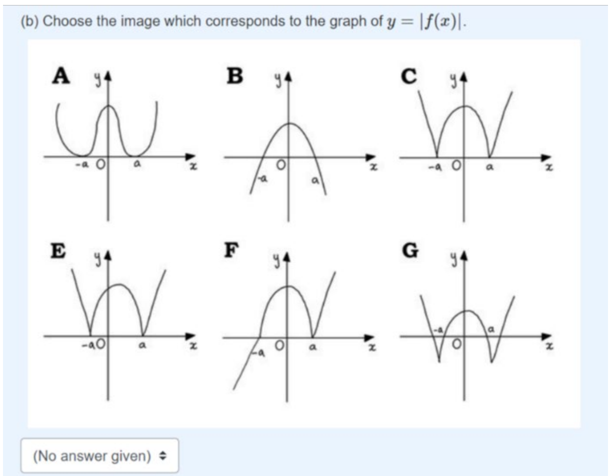


Figure 6: Assessing graphs using a multiple choice question.

Consequently, towards the end of the project, we tried to restructure this question by using JSXGraph. We start by giving the graph of the original function $f(x) = x^2 - a^2$ (Figure 7, left) and by having two sliders the students can modify the graph to create the required one (Figure

7, right). While for this specific question the JSXGraph approach seems adequate, we believe that it will give the desirable outcome only for specific functions and it may put some limitations in the functions the assessors can choose.

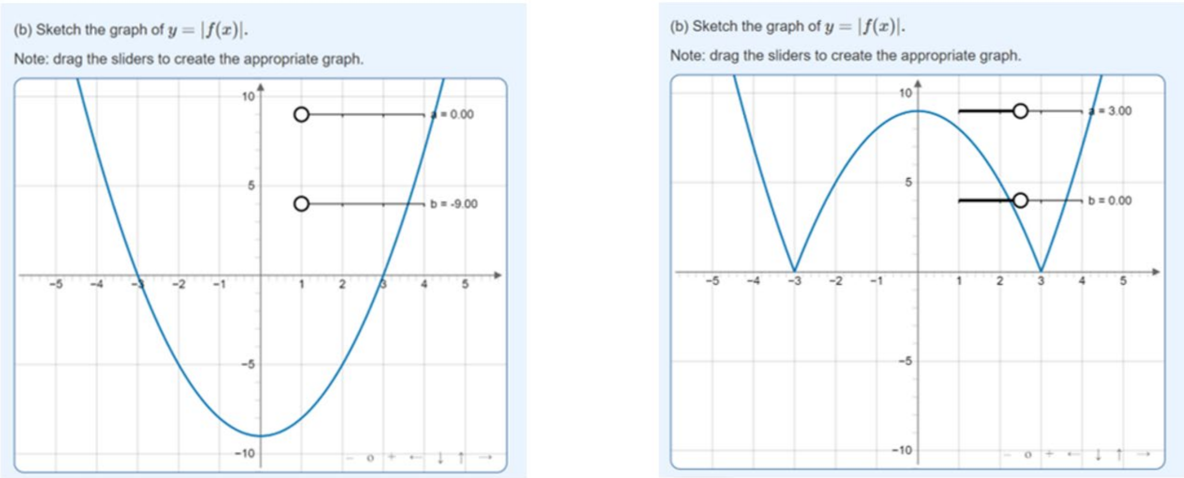


Figure 7: Assessing graphs using JSXGraph. The graph of f(x) = x2 − a2 (left) and the correct answer after moving the two sliders (right).

Effective assessment of “Proof” questions

Assessing ‘Proof’ questions in mathematics, and ‘Derivation’ questions in physics poses inherent challenges due to their written nature and the existence of multiple valid approaches. A possible method for assessment is the ‘fill-in-the-blanks’ model (see Figure 8), offering advantages such as the familiarity of input for students given its similarity to multiple-choice questions (Bickerton & Sangwin, 2022). However, its drawback lies in the highly involved process of constructing questions and the relative ease of completing blanks compared to students working out a proof from scratch.

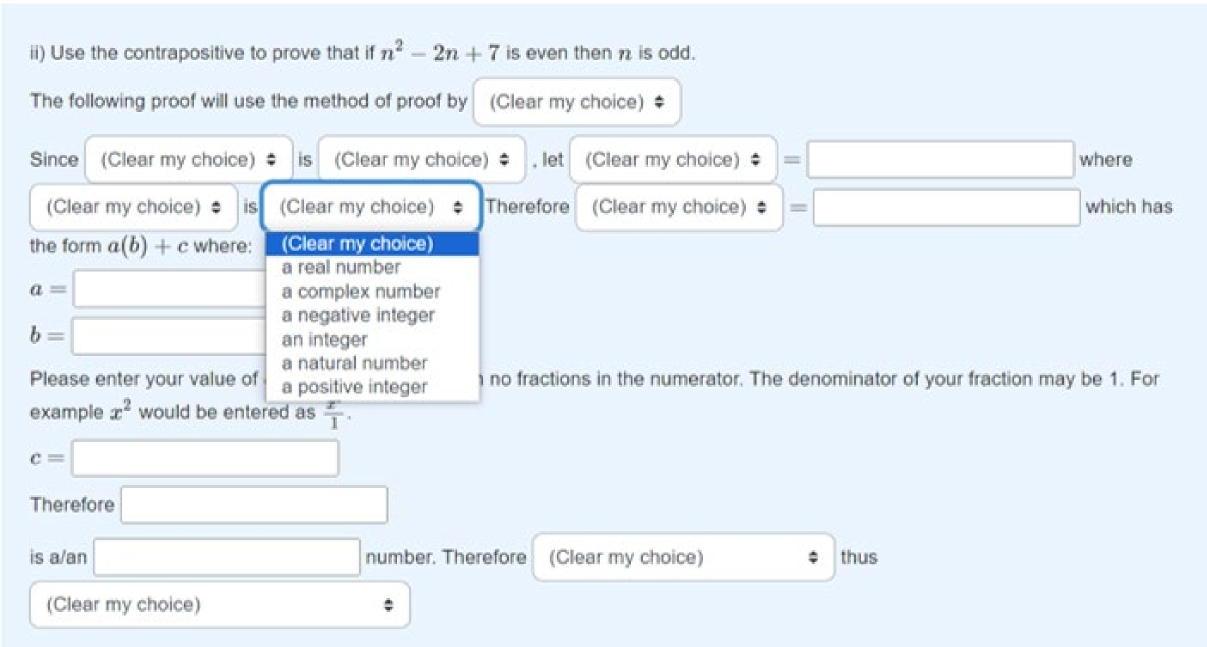


Figure 8: A proof that requires to fill in various inputs.

Another approach is the Parsons proof feature on STACK https://docs.stack-assessment.org/en/Authoring/Parsons/, enabling students to arrange proof steps in the correct order via drag-and-drop. An example of this is given in Figure 9. While more user-friendly and easier to implement than drop-down boxes, it still lacks the complexity required for students to construct proofs from scratch. Hence, while proof assessment is feasible, replicating the same style and level of difficulty remains challenging (Bickerton & Sangwin, 2022).

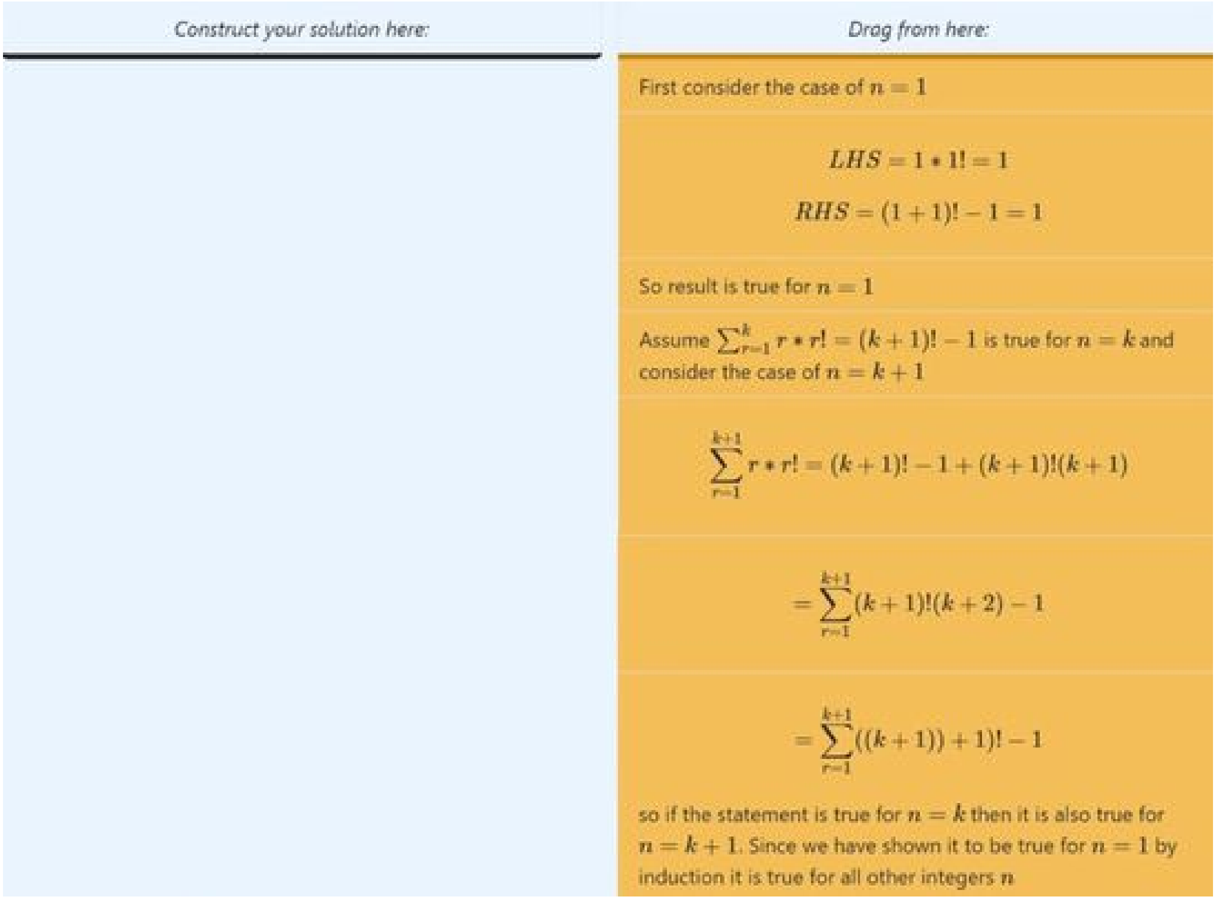


Figure 9: Assessing graphs using a multiple choice question.

Effective Assessment of Short Written Answers

Currently, STACK lacks an in-built feature for grading written answers, as it is primarily designed for STEM subjects where such responses are rare. Despite this, a significant portion of the Physics paper includes questions that require written responses. These questions, as noted in (Burrows et al., 2015), are ideal candidates for Automatic Short Answer Grading systems since they are concise, and focused on assessing content rather than writing style. Furthermore, automatic grading of short answers could potentially reduce bias (Burrows et al., 2015). To handle these types of questions we developed a ‘Proof of Concept’ STACK function. This is an implementation of a regular expression function with the ability to categorise student responses as correct or incorrect by identifying key phrases in a written answer. It incorporates STACK’s built-in Levenshtein distance feature as a basic method of spell-checking, representing just one of many possible approaches to pattern matching and spell-checking within STACK. However, it must be noted that relying solely on this implementation may lead to inaccuracies in categorising all potential student responses, especially when faced with unusual responses from students. Fortunately, with sufficient time and resources, it is possible to develop an automatic short answer grading system that achieves reliability comparable to that of manual marking (Butcher & Jordan, 2010). An example of a question that requires short answer is given in Figure 10.

	(a)	(ii)	Towards the centre of the (horizontal) circle	1	Along the radius (0) Along the radius towards the centre (1)

Figure 10: STACK implementation of a "short answer" question (left) and the SQA official marking scheme (right). The marking scheme is subject to Copyright © Scottish Qualifications Authority.

It is, however, crucial to distinguish these short-answer questions from more extensive essay type questions, such as the ‘Using your Knowledge of Physics’ questions, which are less suited to assessment using this method. Yet, with recent advancements in artificial intelligence, the possibility of accurately automating the assessment of longer written responses is an increasingly achievable one. In 2019, research demonstrated (Hussein et al., 2019) that various models, when trained on data specifically relevant to the content they were assessing, could achieve a correlation of over 90% with human markers. Enhancing these models further would significantly benefit the automated assessment of Advanced Higher Physics.

Awarding partial credit for Differentiation

Manually grading differentiation problems is generally straightforward, yet automating this process presents significant challenges. While basic differentiation questions can be assessed directly, Advanced Higher Mathematics examinations often incorporate more complex problems that involve the product or quotient rules, complicating the allocation of partial credit for correct methodologies. For instance, finding the derivative of a function like

$$f(x) = x^5 \cot(5x)$$

provides an example of this issue as the marking scheme (see Figure 11) rewards students with partial credit for correctly identifying one term in their solution and demonstrating the use of the product rule.

Question			Generic scheme	Illustrative scheme	Max mark
1.	(a)		•[1] evidence of product rule with one term correct [1,4]	•[1] $6x^5 \cot 5x \pm x^6(\ldots)$ OR $-5x^6 \operatorname{cosec}^2 5x + (\ldots)\cot 5x$	2
			•[2] complete differentiation [1,2,3]	•[2] $6x^5 \cot 5x - 5x^6 \operatorname{cosec}^2 5x$	

Figure 11: Official marking scheme for the differentiation question requiring the product rule. The marking scheme is subject to Copyright © Scottish Qualifications Authority.

The foundational knowledge for tackling these types of questions in Advanced Higher Mathematics are the product, quotient, and chain rules. Assuming a student can accurately transcribe a function from one line to another, the correct application of the product rule should yield an answer in a specific format. If a student's response is different from this specified format, it becomes challenging to ascertain whether the product rule was used. This uncertainty introduces difficulties in evaluating such responses, as it is not straightforward to distinguish between the absence of the product rule's application or a mere mistake in the final answer. This predicament highlights the inherent difficulties encountered in grading questions that require the application of the product rule, especially when the expected answer format is not adhered to.

An alternative approach is to use the reasoning by equivalence answer type for the assessment of differentiation questions. This is shown in Figure 12. Equivalence reasoning offers the advantage of allowing students to earn partial credit for correctly executed steps, even if they make an error in arriving at their final answer. This approach, however, requires familiarity with Maxima's functions, particularly the diff() function for inputting answers. The challenge here is that the assessment may inadvertently shift focus from evaluating students' mathematical proficiency to testing their coding skills.

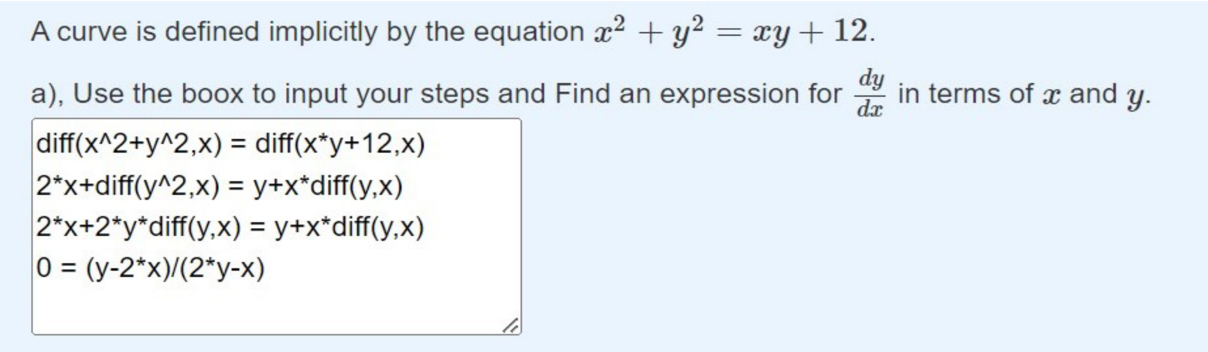


Figure 12: An example of student's answer in a differentiation question using the equivalence reasoning

6. Conclusions

In conclusion, with the exception of open-ended and proof-based questions, utilising STACK for assessing final answers in examinations is highly feasible. The main challenges arise in ensuring compatibility with the SQA's detailed marking criteria, especially regarding partial credit for intermediate steps. However, STACK's rapid developments in recent years, suggests that fully accounting for these marks is within reach.

Currently, about 81% of the Advanced Higher Mathematics curriculum can be automated in STACK, suggesting immediate potential for a partially automated paper alongside some manually graded elements. For Physics, automation has been achieved for approximately 69% of the questions, indicating the need for further research before full automation can be introduced in Scottish schools. However, the automation achieved so far offers the potential for substantial efficiency improvements presenting a strong case for the future automation of Advanced Higher Physics examinations. Therefore, whilst Advanced Higher Mathematics is already well-suited for automation through STACK, some manual grading still remains necessary. Advanced Higher Physics shows promise for complete automation with further research. Given STACK's rapid advancement and our success in aligning existing examination questions with our designed taxonomy, the fully automated assessment of both subjects seems not only feasible but highly likely in the near future.

Acknowlegment:

We would like to thank the Scottish Qualifications Authority for giving us permission to use screenshots of the official marking scheme on this paper.

References

Bickerton, R. T., & Sangwin, C. J. (2022). Practical online assessment of mathematical proof. *International Journal of Mathematical Education in Science and Technology*, *53*(10), 2637–2660. https://doi.org/10.1080/0020739X.2021.1896813

Burrows, S., Gurevych, I., & Stein, B. (2015). The eras and trends of automatic short answer grading. *International Journal of Artificial Intelligence in Education*, *25*. https://doi.org/10.1007/s40593-014-0026-8

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. https://doi.org/10.1016/j.compedu.2010.02.012

Disclosure log. (2023), https://www.sqa.org.uk/sqa/93617.html.

Gratwick, R., Kinnear, G., & Wood, A. K. (2020). An online course promoting wider access to university mathematics. *Proceedings of the British Society for Research into Learning Mathematics*, *40*(1), 1–6. https://bsrlm.org.uk/wp-content/uploads/2020/ 05/BSRLM-CP-40-1-04.pdf

Hussein, M., Hassan, H., & Nassef, M. (2019). Automated language essay scoring systems: A literature review. https://doi.org/10.7287/peerj.preprints.27715v1

Kinnear, G., Wood, A. K., & Gratwick, R. (2021). Designing and evaluating an online course to support transition to university mathematics. *International Journal of Mathematical Education in Science and Technology*, *0*(0), 1–24. https://doi.org/10.1080/ 0020739X.2021.1962554

Kinnear, G. (2018). Delivering an online course using STACK. *Proceedings of the STACK Conference*. https://doi.org/10.5281/zenodo.2565969

Orthaber, M., Stutz, D., Antretter, T., & Ebner, M. (2020). Concepts for e-assessments in¨ stem on the example of engineering mechanics. *International Journal of Emerging Technologies in Learning (iJET)*, *15*(12), pp. 136–152. https://doi.org/10.3991/ijet. v15i12.13725

Sangwin, C. J., & Kocher, N. (2016). Automation of mathematics examinations. Computers and Education, 94, 215–227. https://doi.org/10.1016/j.compedu.2015.11.014

Sangwin, C. J., & Zerva, K. (2020). Developing online learning materials to support undergraduate education at the University of Edinburgh. *Mathematics Today*, *56*(3), 98– 101.

Sangwin, C. (2019). Developing and evaluating an online linear algebra examination for university mathematics. *Eleventh Congress of the European Society for Research in Mathematics Education*, (15).

SQA 2019 advanced higher mathematics finalised marking instructions, (2019). https://www.sqa.org.uk/pastpapers/papers/instructions/2019/mi_NAH_Mathematics_mi_2019.pdf

SQA 2019 advanced higher mathematics past paper, (2019). https://www.sqa.org.uk/pastpapers/papers/papers/2019/NAH_Mathematics_QP_2019.pdf

SQA 2019 advanced higher physics finalised marking instructions, (2019). https://www.sqa.org.uk/pastpapers/papers/instructions/2019/mi_NAH_Physics_mi_2019.pdf

SQA 2019 advanced higher physics past paper, (2019). https://www.sqa.org.uk/pastpapers/papers/papers/2019/NAH_Physics_all_2019.pdf

Stephen, J. P. (2022). The SQA official statistics attainment statistics.

Thake, A. (2023, October). Making it click: The case for digital examinations in England.

Biographical notes

Konstantina Zerva studied Physics at University of Ioannina, Greece and I received her PhD in 2013 from the same University. After that, she worked as a Teaching Assistant at the School of Physics and Astronomy and at the School of Mathematics at the University of Edinburgh.

Since 2017 she has been working as Facilitator in Technology Enhanced Mathematical Education at the School of Mathematics at the University of Edinburgh. Her main role is to design and develop online assessments for the undergraduate Mathematics courses using the STACK assessment system and also to give practical support to colleagues about the best use of e-assessment. She has a growing interest in scholarship of learning and teaching and how technology and e-learning can be used to support learners and educators. In 2019 she was awarded a fellowship in the Higher Education Academy.

Ben Chooyin is a final year student at the University of Edinburgh and is studying for a degree in Mathematics (BSc Hons).

Jack Fortune is a final year student at the University of Edinburgh and is studying for a degree in Mathematics and Physics (BSc Hons).

Rachel Hollingsworth is a final year student at the University of Edinburgh and is studying for a degree in Mathematics and Physics (BSc Hons).

Mitchell McLachlan is a final year student at the University of Edinburgh and is studying for a degree in Mathematics (BSc Hons).

Heather Napthine is a final year student at the University of Edinburgh and is studying for a degree in Mathematics (BSc Hons).

Focus: Other topics related to STACK.

Article number: 09

New Developments Around the Evaluation Tool STACKrate

Daniel Meißner*; Ruhr-Universität Bochum, Germany

Jonas Lache; Hochschule Ruhr West and Ruhr-Universität Bochum, Germany

Abstract

This paper highlights recent developments in STACKrate, a JavaScript-based tool for the evaluation of STACK questions using a star rating principle. We will discuss the new features and enhancements that have been added since the initial release of STACKrate in 2022 and which improve the functionality for educators. We will also present the user-friendly STACKrate Snippet Generator, which simplifies the process of integrating STACKrate into STACK questions. Finally, we will outline the planned adaptation of STACKrate to the new STACK-JS functionality to ensure compatibility with future versions of STACK.

* Corresponding author: daniel.meissner-i4k@ruhr-unibochum.de

1. Introduction

The principle of star ratings is a simple and common way to assess a level of satisfaction or agreement, or to express the degree of expression of a characteristic. Star ratings are ubiquitous in everyday life. For example, they are used to assess customer satisfaction in restaurants (Mathayomchan & Taecharungroj, 2020), for product reviews (Dong et al., 2020), for nutrition labels on product packaging (Shahid et al., 2020), and as a measure of hospital quality (Bilimoria & Barnard, 2021). The tool “STACKrate”, a free and open-source JavaScript tool, uses star ratings to evaluate STACK questions (Lache & Meißner, 2022). STACK is a question type for creating digital mathematics tasks within the learning management system Moodle (for more information on STACK, see Sangwin, 2015). In order to find out how students perceive STACK questions, an evaluation of the questions can be helpful. After a brief introduction to STACKrate and its features, this paper outlines the new features of the tool in the recently released version v0.2, as well as the newly developed “STACKrate Snippet Generator”, which simplifies the use of STACKrate for task creators. The paper concludes with a discussion of future challenges and perspectives.

2. The evaluation tool STACKrate

STACKrate allows students to rate a STACK question by choosing between one and five stars. On the one hand, this rating system offers a convenient and efficient user experience for students. In the best scenarios, they simply need to click twice to submit their rating. On the other hand, task creators can effortlessly integrate a series of STACKrate evaluation questions into a STACK task by copying and pasting code snippets into the question or feedback text. Rating results are written to a hidden STACK input field in JSON format and can be exported via Moodle quiz reports. Notably, no additional software or external database is required. STACKrate evaluation boxes can be customised in a number of ways. Firstly, a STACK question can contain an unlimited number of STACKrate evaluation questions. The appearance of the box can be easily customized using CSS: for example, the text colour, background colour and border of the box can be changed to suit personal preferences. STACKrate evaluation boxes can also be expanded to include a free comment area where students can explain their ratings and provide additional details. This feature has proven to be helpful in understanding the challenges students encounter during their engagement with a STACK task and in identifying bugs (see Lache & Meißner, 2022). Once the students answered all rating questions, the Moodle “Check” button will change its name to “Submit ratings”. Finally, STACKrate has been translated into several languages, including German, Spanish, French, Dutch, and Portuguese. Overall, a STACKrate evaluation looks as shown in Figure 1.

Experience has shown that STACKrate can generate high response rates (Lache & Meißner, 2022). Since the first release of STACKrate in 2022, the tool has been used at several universities (e.g. by Eichler et al., 2022). The user feedback we have received has been very helpful in identifying bugs to be fixed and sensible features to be added in an updated version of STACKrate. These changes are described in the following section.

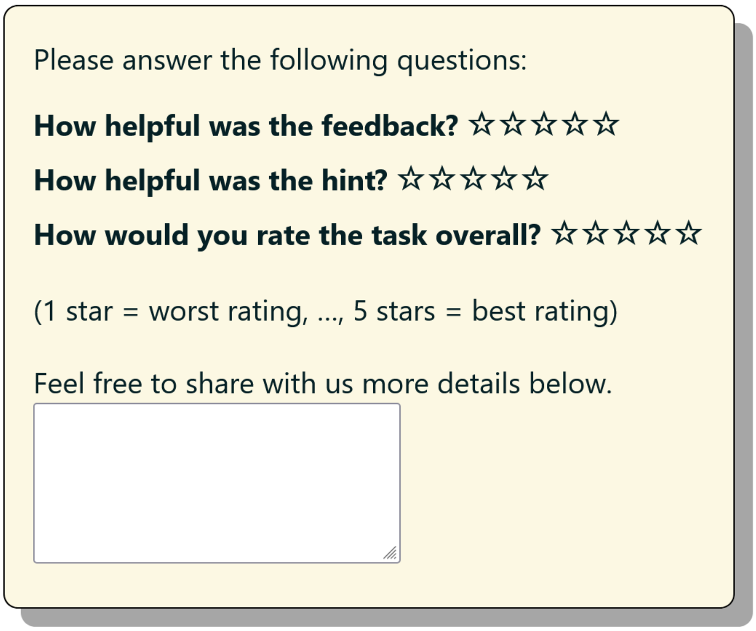


Figure 1: Screenshot of a STACKrate evaluation box with text area for comments.

3. New features in STACKrate v0.2

The updated version of STACKrate was released in spring 2024 and includes several new features and bug fixes. These are designed to meet the needs of the users and make the tool more flexible.

As of STACKrate v0.2, users can choose a custom highlight colour for the stars. When students hover over the stars or click on a star to select a certain number of stars, the colour of the stars changes. Previously, the star highlight colour was hardcoded to orange, which led to bad contrast depending on the background colour chosen for the STACKrate evaluation box (see Figure 2, top). The newly added possibility to use a custom highlight colour solves this problem (see Figure 2, bottom, where a dark shade of blue is used for better contrast).

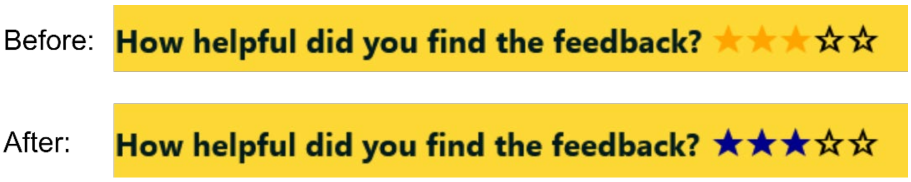


Figure 2: Custom highlight colour for the stars.

With regard to the highlighting of the stars, STACKrate v0.2 also includes a bug fix: When students select and then deselect stars, the stars are first highlighted (as previously mentioned) and are then reset to their initial colour. However, when task creators selected a custom text colour for their evaluation box, the stars' colour did not change back to the custom text colour, but always to black (see Figure 3, top). After fixing this, the stars change back to the custom colour after highlighting (see Figure 3, bottom).

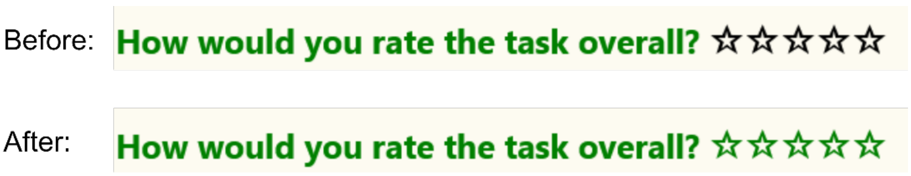


Figure 3: Bug fix regarding the star colour after highlighting.

Perhaps the biggest adjustment in STACKrate is the change in the default behaviour of whether students have to submit all ratings. In the previous version of STACKrate, students could only submit the quiz once all the STACKrate evaluations had been submitted, and task creators had no way of changing this. Due to popular demand, the default behaviour is no longer to require students to submit all their ratings in order to submit the quiz. However, if the former behaviour is still desired, it can be invoked by setting an option for the STACKrate evaluation box.

Another feature that has been requested is a possibility to disable the hardcoded legend (in English it is “1 star = worst rating, …, 5 stars = best rating”). Although the auto-generated legend simplifies the setup of a STACKrate evaluation box and is a valuable feature, users may want to use no legend or a custom legend. Therefore, an option to disable the default legend was a reasonable request and has been implemented.

Finally, STACKrate has been adapted to new conditions in Moodle and is now fully compatible with Moodle 4. The updated documentation provides further details about the new features and bug fixes of STACKrate: https://www.ruhr-uni-bochum.de/stackrate-maths/docs

4. STACKrate Snippet Generator

Creating the markup for a STACKrate rating form can pose a significant challenge, particularly for newcomers to HTML. However, even experienced users may struggle with writing error-free HTML markup, such as ensuring all appropriate end tags are included. To simplify this process, STACKrate’s website now features a user-friendly graphical interface known as the STACKrate Snippet Generator. This tool enables users to effortlessly customize the description and questions of their rating form using a What You See Is What You Get (WYSIWYG) approach (see Figure 4). This enhances the accessibility for all users, especially for those with limited HTML experience.

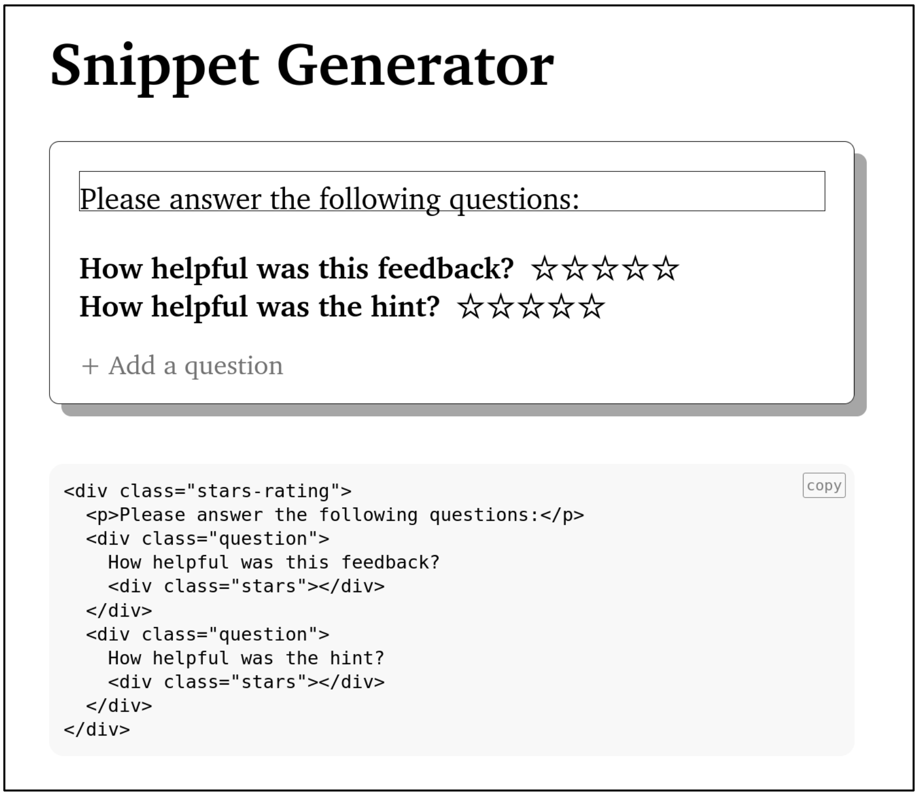


Figure 4: The STACKrate Snippet Generator.

As users make adjustments to their rating form within the interface, the Snippet Generator dynamically updates the HTML markup displayed below the form. By utilizing the copy button, users can easily transfer the generated markup to their clipboard for seamless integration into a STACK question. Presently, users have the flexibility to modify the description and rating questions, including the possibility to add or remove questions as needed. In the pipeline are plans to introduce additional functionalities, such as the ability to edit the legend and incorporate a comment box for free-form comments. This ongoing development aims to further streamline the process of creating and customizing rating forms with STACKrate.

The evolving STACKrate Snippet Generator can be accessed at https://www.ruhr-uni-bochum.de/stackrate-maths/generator. Feedback, whether in the form of comments or feature requests, is welcomed to continually improve the tool.

5. Challenge: Ban of script elements in STACK questions

Allowing users to embed arbitrary script elements into STACK questions poses a significant security risk. While we may trust teachers and question authors not to misuse JavaScript, it complicates the sharing of materials as users must carefully review shared content for potential malware. Since question authors are not security experts, expecting them to detect and prevent malicious scripts is unreasonable. JavaScript can be exploited to interact with the Moodle site under the user's login, enabling actions like submitting quiz responses or posting messages on course forums on behalf of the user. To address this security concern, there are plans for STACK to prohibit the insecure use of script elements (see section "The general security reason" at STACK Docs, n.d.).

However, many question authors and developers have grown accustomed to the inclusion of HTML script elements and have utilized them for beneficial purposes. One such example is STACKrate. Consequently, rather than outright banning script elements, STACK aims to empower users by offering a secure means of implementing extensions. The proposed solution is STACK-JS, which utilizes iframe elements isolated from the virtual learning environment (VLE). Within these iframes, a JavaScript library facilitates secure communication with the STACK question, ensuring that the script's impact is confined to the specific question (STACK Docs, n.d.).

The shift towards prioritizing STACK-JS over script elements also impacts STACKrate, which relies on script elements to function. Specifically, STACKrate saves its rating outcomes in a concealed STACK input field, necessitating access to it. Fortunately, the design of STACK-JS caters to this requirement and offers the function `request_access_to_input`, enabling connectivity to a STACK input field from within the iframe (STACK Docs, n.d.). However, beyond linking to a hidden input field, additional access is needed to modify the check button to display 'submit rating'.

Another hurdle is ensuring that new iterations of STACKrate remain compatible with STACK releases that do not restrict the use of script elements but lack support for STACK-JS. A seamless solution is sought to provide users with a consistent experience, regardless of the STACK version utilized on their Moodle site. This implies that the STACKrate script detects whether it is running inside of an iframe element or not and adapting its behaviour respectively.

6. Conclusion

With the recent release of version v0.2, STACKrate has undergone significant improvements, bug fixes and introducing new features that enhance the customisation options for users. This includes changing the highlight colour, customizing the legend and adding support for Moodle 4. The addition of the Snippet Generator has greatly simplified the creating of rating forms, making STACKrate more user-friendly for a broader audience. Detailed documentation with screenshots and video tutorials on the website enable users to swiftly initiate the process of having their students evaluate their STACK questions. Users are encouraged to explore the Snippet Generator and share thoughts to help enhance its functionality and user experience. Valuable feedback, also concerning STACKrate as a whole, will contribute to making it a useful and effective tool to evaluate STACK questions.

More information on STACKrate, the documentation and the Snippet Generator can be found at https://www.ruhr-uni-bochum.de/stackrate-maths.

Bibliography

Bilimoria, K. Y., & Barnard, C. (2021). An Evolving Hospital Quality Star Rating System From CMS: Aligning the Stars. *JAMA, 325*(21), 2151. https://doi.org/10.1001/jama.2021.6946

Dong, J., Chen, Y., Gu, A., Chen, J., Li, L., Chen, Q., Li, S., & Xun, Q. (2020). Potential Trend for Online Shopping Data Based on the Linear Regression and Sentiment Analysis. *Mathematical Problems in Engineering*, 2020, 1–11. https://doi.org/10.1155/2020/4591260

Eichler, A., Liebendörfer, M., Schürmann, M., & Sommerlade, F. (2022). Developing digital networks for learning and teaching mathematics in introductory courses. *Fourth conference of the International Network for Didactic Research in University Mathematics*, Hannover.

Lache, J. A., & Meißner, D. (2022). How to STACKrate: The Development of an Intuitive Evaluation Tool. *International Journal of Emerging Technologies in Learning (iJET), 17*(23), 7–14. https://doi.org/10.3991/ijet.v17i23.36527

Mathayomchan, B., & Taecharungroj, V. (2020). “How was your meal?” Examining customer experience using Google maps reviews. *International Journal of Hospitality Management, 90*, 102641. https://doi.org/10.1016/j.ijhm.2020.102641

Sangwin, C. (2015). Computer Aided Assessment of Mathematics Using STACK. In *Selected Regular Lectures from the 12th International Congress on Mathematical Education* (pp. 695–713). Springer International Publishing. https://doi.org/10.1007/978-3-319-17187-6_39

Shahid, M., Neal, B., & Jones, A. (2020). Uptake of Australia’s Health Star Rating System 2014–2019. *Nutrients, 12*(6), 1791. https://doi.org/10.3390/nu12061791

STACK Docs. (n.d.). STACK-JS [Documentation]. *STACK Docs*. Retrieved 31 May 2024, from https://docs.stack-assessment.org/en/Developer/STACK-JS/

Information on the authors

Dr. Daniel Meißner

Daniel is a research assistant at Ruhr-University Bochum with a PhD in mathematical physics. Lately, he is actively involved in two OERContent.nrw projects aimed at creating open educational resources for probability theory and statistics.

Jonas Lache

Jonas is a research assistant at the Ruhr-Universität Bochum and a PhD student in mathematics education. He is also a member of the e-learning team at the Hochschule Ruhr West, where he is responsible for the Moodle question type STACK.

Focus: STACK in teaching or exams.

Article number: 10

On ongoing extensions of a comprehensive collection of mathematical STACK problems in German language

Maike Schelhorn*, Oleg Boruch Ioffe, Katharina Jaenicke, Reik V. Donner

Magdeburg-Stendal University of Applied Sciences

Abstract

Since the winter term 2022/23, Magdeburg-Stendal University of Applied Sciences has been using STACK problems primarily in the mathematics courses for civil engineers. Most of the problems initially selected have been taken from two existing, partially overlapping databases with mathematical STACK problems in German language, which have been collected by the Ruhr University of Bochum (‘DOMAIN’) and the Cologne University of Applied Sciences (‘Digitaler Aufgabenpool Mathematik’). We have systematically categorized the existing problems according to the topics covered by our local courses, and addressed a few issues with wording, notation or technical implementation whenever identified in the initial screening phase or during their practical use in our regular courses.

During our introduction of STACK problems in our courses, we have identified several topics in our mathematics for engineering curricula, which have not yet been well covered by the existing problem collections in German language. Accordingly, we have started systematically addressing such underrepresented topical fields. While some areas are currently being filled with translations of existing problems in English language, others require the development of new STACK problems to fill the present gaps. Our aim is to provide our students with a comprehensive collection of digital problems covering all topics of our mathematics courses in the first two study years.

* Corresponding author: maike.schelhorn@h2.de

1. Introduction

The use of STACK problems in higher education opens up many possibilities such as automatic feedback based on complex feedback trees, randomisation of variables and the use of (interactive) graphics. As part of the project 'h²d² - didaktisch und digital kompetent Lehren und Lernen' at Magdeburg-Stendal University of Applied Sciences, which is financially supported by the 'Stiftung Innovation in der Hochschullehre', the decision was made to switch from the previously used digital problems employing the commercial tool WIRIS to STACK (Judakova et al., 2023).

STACK problems were first introduced into the 'Mathematics 1' and partially also 'Mathematics 3' courses for civil engineers in the winter term 2022/23. Since the summer term 2023, they were also integrated into the 'Mathematics 2' course of the same study program and have since become an integral part of all three mathematics courses for civil engineers. Meanwhile, STACK problems are also being used in the mathematics courses for the bachelor curricula on recycling and waste management, water management (all held in German language), and the study programme on sustainable resources economics and management ('STREaM') taught in English. Moreover, STACK problems for many different areas of mathematics are available to all students in a Digital Mathematics Learning Support Centre. In the different curricular courses, the STACK problems are used both for weekly self-assessments and for voluntary e-assessments held every three to four weeks in order to obtain bonus points for the exam.

2. A German collection of digital STACK problems and their adaptation

For the described purposes, a large quantity of well-developed STACK problems covering the variety of different topics of engineering mathematics were required. In order to be able to offer our students a sufficiently large collection of problems, existing databases were used as a starting point for further adaptations and extensions. Specifically, we have selected material from 'DOMAIN' (https://db.ak-mathe-digital.de, Ruhr University Bochum) and 'DIGITALER AUFGABENPOOL MATHEMATIK' (https://aufgabenpool.th-koeln.de, Cologne University of Applied Sciences).

The existing problems have been categorised according to the curricula of our mathematics courses for civil engineers, scanned for duplicates and tested for correct functionality. Afterwards all materials were checked for technical and didactic issues. For example, specific attention was paid to the use of understandable language, suitable notation for our courses and feedback in form of a detailed sample solution. The materials are being continuously revised after each course based on the recorded test attempts (within the learning management platform moodle) and other feedback from students and teachers, including the results of structured questionnaires, qualitative interviews with selected students performed by collaborating social scientists (Donner et al., 2023), and suggestions directly communicated to the teachers by individual students. A more detailed description of our initial collection has already been provided in Judakova et al. (2023).

3. Extensions of the collection

Even though the described collection of mathematical STACK problems in German language already contained a large amount of individual problems, not all the topics taught in our current curricular mathematics classes were fully covered. On one hand, there exist subject areas in our curricula for which there were no or not enough suitable STACK problems available. On the other hand, to create self-study and learning support materials specifically in the domain of statistics and probability theory (which are relevant beyond our engineering faculties, especially in the field of social sciences and economics), many topically suitable problems from this field were required, while those have been hardly covered previously in the existing databases. In order to be able to deal with these issues, it was necessary to further expand our initial collection. This extension process is currently ongoing, exploiting German translations from existing English STACK problems along with the development and implementation of new problems by our team.

Translation of STACK problems based on the HELM workbooks

Motivated by the ongoing development of a digital self-study course in the field of statistics, we have started to translate existing problems from a pool of STACK problems in English language. Specifically, we have based our attempts on previous efforts undertaken by teams from the University of Edinburgh and the Loughborough University who developed STACK problems based on the HELM (Helping Engineers Learn Mathematics) workbooks (Zerva et al., 2022).

In addition to the plain STACK problems, the HELM pool also contains information texts and tasks of question types other than STACK. For this reason, the materials are very suitable for the use in a self-study course. The intention is to give our students the opportunity to study or repeat the basics of statistics on their own. Students can then complete the course independently at the time when they need the knowledge, for example when working on their final thesis. We have selected the following topical areas from the HELM pool and then translated the corresponding materials into German language: Sets and Probability, Descriptive Statistics, Discrete Probability Distributions, Continuous Probability Distributions and Normal Distributions. Other existing sections of the HELM pool could be translated into German as well as part of our future work given sufficient demand by our students and/or teachers.

Although the materials mentioned above have already been translated, at this point there are still some challenges to be addressed. For example, most common German notations used in statistics and probability theory may partially differ from those used in HELM. In some cases, there even exist different parallel notations describing the same mathematical entity. In such cases, a careful selection needs to be made in order to guarantee maximal coherency between the STACK problems, German textbooks and the notations used by our teaching staff and hence provide our students with a self-study course that they can understand. To support this step, we are currently initiating an internal stakeholder dialogue with the teaching staff of mathematics and statistics at all five faculties of the Magdeburg-Stendal University of Applied Sciences to find a common solution for this issue.

Development of new STACK problems

At present, we are also systematically identifying any missing or underrepresented topics that are part of our mathematics for engineering curricula. All these topics are summarized and then prioritized. The assigned individual priority depends on when the topic is covered in the courses, whether or not at least a few suitable STACK problems already exist, how many problems are needed to fully cover that topic, and how complex the development of the material is expected to be. STACK problems for the highest prioritized topics are being developed and implemented first. Examples of gaps that have already been addressed by new developments include the solution sets of inequalities, trigonometric equations, and the application of l'Hôpital's rule.

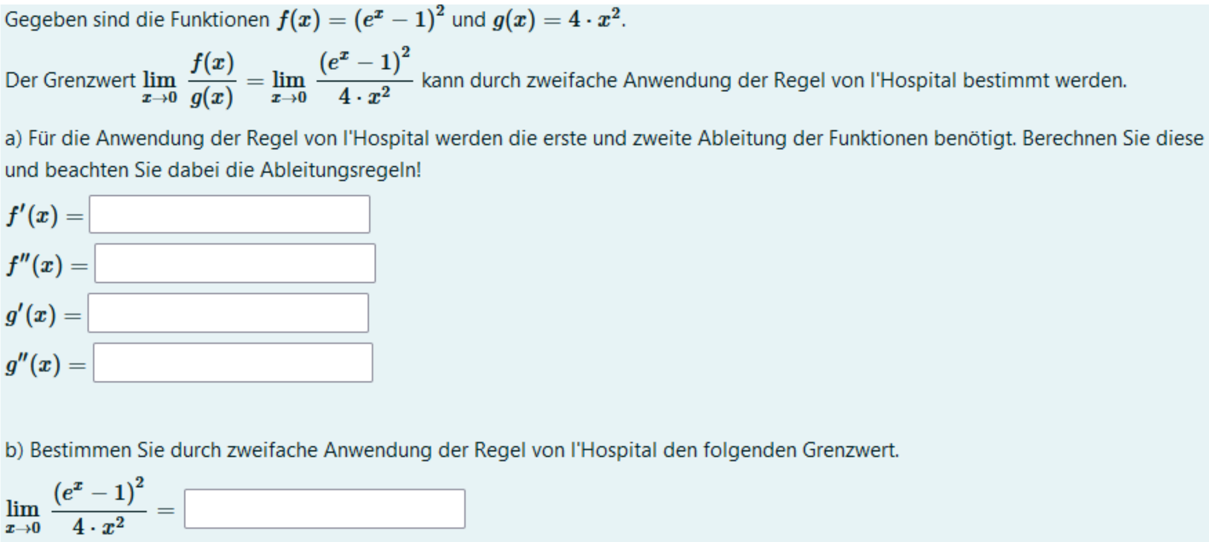


Figure 1: Example of a newly developed STACK problem on the twofold application of l'Hôpital's rule

Figure 1 shows a designed problem for l'Hôpital's rule, which is used for the evaluation of limit values for mathematically indeterminate expressions. In this example, a twofold application of the rule is required. This results in a higher level of requirements than a single application, which is the reason for the problem being divided into two parts. By this division, we expect to foster the student's attempts in completing the set task. In the first part of the problem, students are asked to calculate all necessary derivatives. In the second part, the limit value is then to be determined by applying the rule twice. The problem is intended for students who are already familiar with the single one-time application of the rule and should therefore not be used as an introductory problem.

To illustrate a given mathematical topic or offer other approaches to coping with it, interactive graphics using JSXGraph (https://jsxgraph.uni-bayreuth.de/wp/index.html) are being used where appropriate. For example, we employed JSXGraph in newly developed STACK problems in the field of trigonometric equations. Figure 2 shows a new problem consisting again of two parts, which offers a visual interpretation of the solutions of a trigonometric equation by using JSXGraph in the problem statement. In the first part, students are informed that the solutions of the given trigonometric equation are the x-values of the intersection points between a constant and a trigonometric function. They are then asked to use the interactive graphic to display the given equation. In the second part, the students are requested to find the solutions to the equation in a given interval. The STACK problem allows students to explore trigonometric equations in a visual way. It was designed as a starting point into the topic of trigonometric equations, so students are not asked to analytically calculate the corresponding solutions in this case.

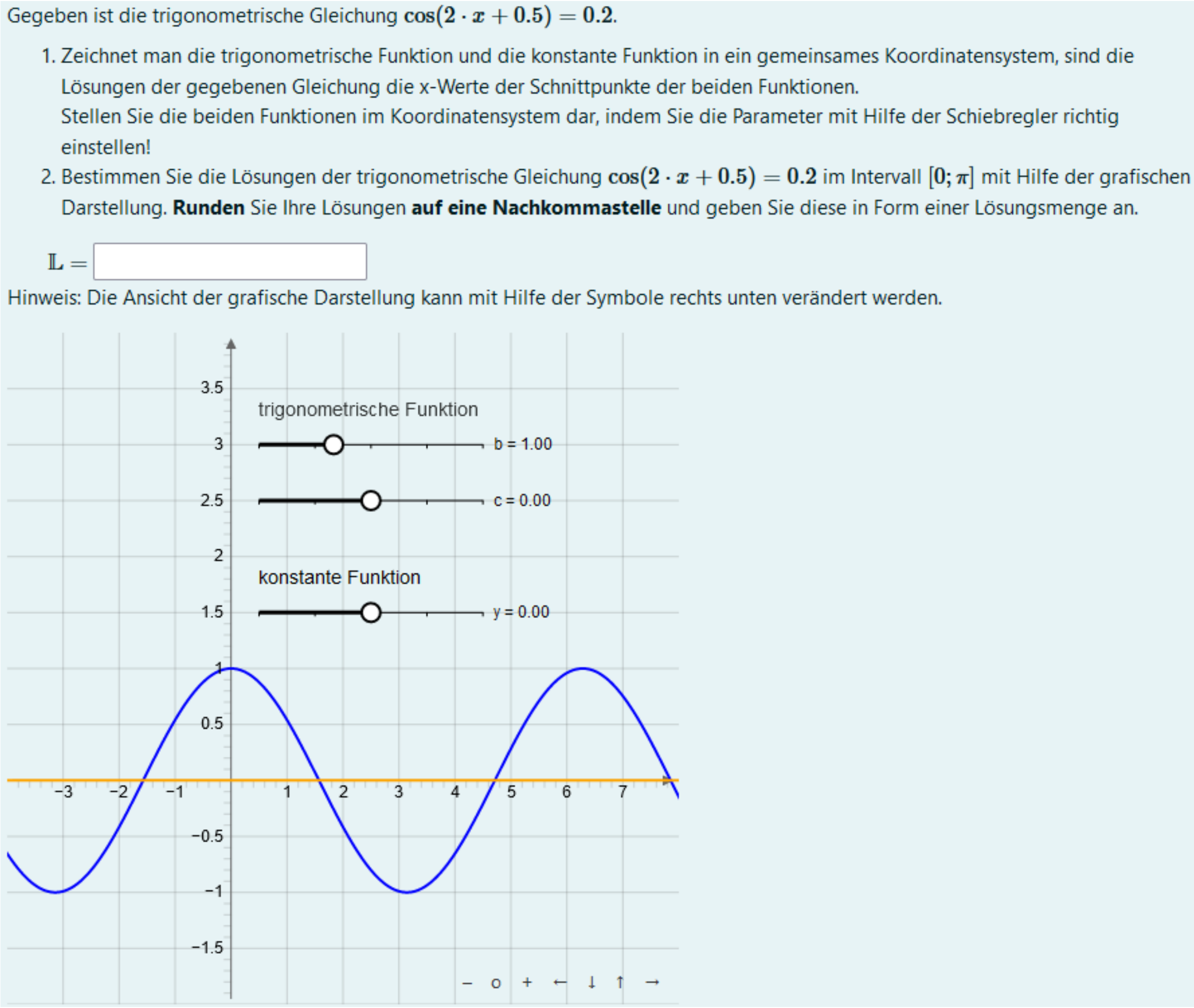


Figure 2: Example of a newly developed STACK problem allowing to visualise and graphically solve trigonometric equation

4. Outlook on future work

Our attempt to develop a new comprehensive collection of mathematical STACK problems in German language is still in progress: There are still several gaps of underrepresented topics that need to be filled with newly developed STACK problems. Moreover, we are dealing with the described issues of the translation of existing English problems (especially on probability theory and statistics) from the HELM database concerning a unifying notation meeting German standards.

While our current focus is still on providing a full topical coverage of the different engineering mathematics courses, it might be necessary to further expand the collection due to the different needs of other mathematics and statistics courses at the Magdeburg-Stendal University of Applied Sciences. In this context, further translations of HELM problems could be carried out for other mathematical fields. Moreover, additional translations of the new German STACK problems developed by our team into English may become necessary to also provide a uniform and as complete as possible topical coverage of existing German and English course programs at our university.

In any case, we plan to publicly release our extended collection of mathematical STACK problems in German language as open educational resource by the end of 2025 at the latest.

Bibliography

Donner, R. V., Judakova, G., Ioffe, O. B., & König, L. (2023). *Integration digitaler Mathe-matik-Aufgaben in die ingenieurwissenschaftliche Grundlagenausbildung*. In H. Dölling, C. Schäfle, S. Kürsten, M. Hunger, J. Hirtt, & P. Riegler (eds.), *Tagungsband zum 5. Symposium zur Hochschullehre in den MINT-Fächern* (pp. 214-221). BayZiel. https://doi.org/10.57825/repo_in-4434

Judakova, G., Ioffe, O. B., König, L., & Donner, R. V. (2023*). A Digital Mathematics Leaning Support Centre Based on a Curated German-Language Collection of Mathematical STACK Problems*. *International Meeting of the STACK Community 2023*, Tallinn, Estonia. https://doi.org/10.5281/zenodo.8032271

Zerva, K., Sangwin, C., Jones, I., & Quinn, D. (2022). *Rejuvenating the HELM Workbooks as Online STACK Quizzes in 2020*. International Journal of Emerging Technologies in Learning (iJET), 17(23), pp. 69–88. https://doi.org/10.3991/ijet.v17i23.35923

Information on the authors

Maike Schelhorn

Research Associate in the MaSta Lab of the project ‘h²d² - didaktisch und digital kompetent Lehren und Lernen’ at Magdeburg-Stendal University of Applied Sciences since 2023

Oleg Boruch Ioffe

Research Associate in the MaSta Lab of the project ‘h²d² - didaktisch und digital kompetent Lehren und Lernen’ at Magdeburg-Stendal University of Applied Sciences since 2021

Katharina Jaenicke

Student Assistant in the MaSta Lab of the project ‘h²d² - didaktisch und digital kompetent Lehren und Lernen’ at Magdeburg-Stendal University of Applied Sciences since 2021

Reik V. Donner

Professor of Mathematics (especially Stochastic Modeling and Data Science) at Magdeburg-Stendal University of Applied Sciences since 2018, head of the MaSta Lab since 2021

Focus: STACK in teaching or exams.

Article number: 11

On the Testing of Linear Algebra with STACK System

Masumi KAMEDA*; Freelance,

Mitsuru UDAGAWA; Nagoya University

Abstract

This report examines the implementation of Computer-Based Testing (CBT), specifically Internet-Based Testing (IBT), in a higher education setting for a Linear Algebra course. The Learning Management System (LMS) used was Moodle, with web-based quizzes created using the STACK mathematics online assessment system. Initially, the methodology involved 15 face-to-face classes, culminating in a final class where all students gathered in a university classroom to take a web-based test under direct supervision, using their personal laptops. This approach was utilized for approximately five years. However, during the COVID-19 pandemic in 2020, the format transitioned to 15 remote classes, followed by a final in-person class in the university classroom where students completed their exams on laptops. In the academic years 2021 and 2022, after conducting 15 remote sessions, students took their exams via web tests at individual learning locations. This report details the practices associated with these IBT implementations.

* Corresponding author: mxcnn668@yahoo.co.jp

1 ICT, and DX

In this section, the authors interpret the keywords necessary for evolutionary operations in the era of the New Normal.

Information and Communication Technology (ICT)

ICT refers to communication using technology. It is a generic term for not only information processing but also industries and services that use communication technologies such as the Internet.

Digitization

Digitization involves converting processes to digital, such as transforming paper-based customer lists into databases and automating tasks like copying and pasting with RPA, to enhance business efficiency and reduce costs. Figure 1 is a photograph of students collectively learning using their own mobile devices and laptops.

Fig. 1: A photograph of students collectively learning

Fig. 2: A photograph of students individualized learning

Digitalization

Using digital technology to transform business models creates new value and customer experiences, such as shifting from owning cars to car-sharing and from DVD rentals to streaming services. Figure 2 is a photograph of students individualized learning through their own smartphones.

Digital Transformation (DX)

New digital entrants are disrupting industries, urging companies to embrace digital transformation, but challenges like outdated systems and resource constraints persist. Figure 3 is a photograph students learning through their mobile laptops during remote lectures. Figure 4 is a photograph of learning individually and in groups through screen-based interactions.

Fig. 3: A photograph of student learning during remote lecture

Fig. 4: A screenshot of learning through screen-based interactions

2 EduTech

In this section, the authors explain the educational system they have used in higher education in an evolutionary way. In other words, the authors will explain educational techniques (EduTech).

Face–to–face learning

From 1987 to 2012, this era marked the integration of personal technology in education, where students used their mobile laptops for learning in a traditional classroom setting.

Blended learning

From 2013 to 2019, the e-Learning system was introduced to complement face-to-face learning, creating a blended learning environment that combined online educational materials and opportunities for interaction online with traditional place-based classroom methods.

Distance learning

From 2020 to 2022, the advent of web meeting systems was added to the blended learning model, facilitating distance learning. This stage represents a shift towards fully online classes and remote education, necessitated by global circumstances and enabled by technological advancements. Figure 5 is the knowledge delivery by teachers through distance learning. Figure 6 is the screenshot of taking a math test with random numbers in a distance learning class.

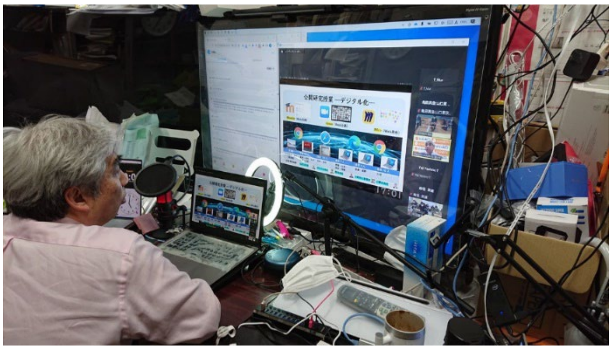


Fig. 5: Knowledge delivery by teachers through distance learning

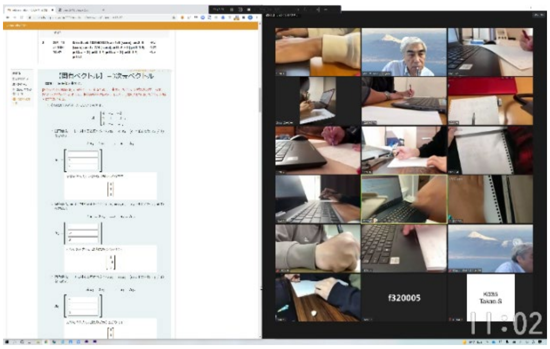


Fig. 6: A view of taking a math test with random numbers in a distance learning class

3 Educatinal environments

In this section, the authors explain the overview of the e-Learning Client/Server System, e-Learning computing system, and learning and teaching environment.

Client/Server system

Client system: Students utilize their own mobile laptops and smartphones, adopting a Bring Your Own Device (BYOD) approach. This flexibility allows learners to access educational resources anytime, anywhere, fostering a more personalized and convenient learning environment. Figure 7 is the learning by desktop computers installed by the university, and own mobile laptop devices of students (Practical training for each student).

Server system: As teachers and system engineers, we provide a centralized server system functioning as a Learning Management System (LMS). This platform supports the administration, documentation, tracking, reporting, and delivery of educational courses, training programs, or learning and development programs. Figure 8 is the conceptual diagram for e-Learning system (Internet layer, intranet layer, and deeper intranet layer).

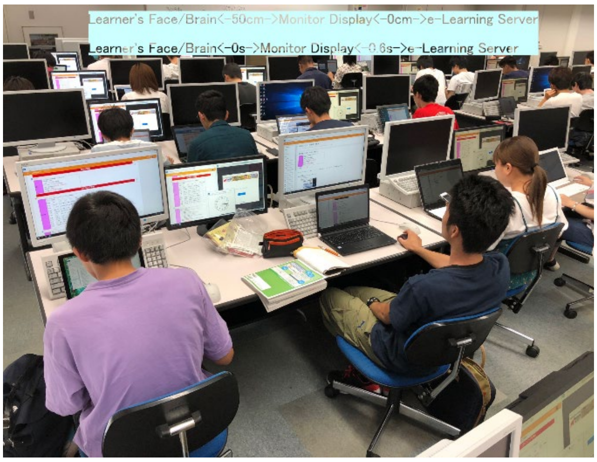


Fig. 7: A photograph of the learning by desktop computers

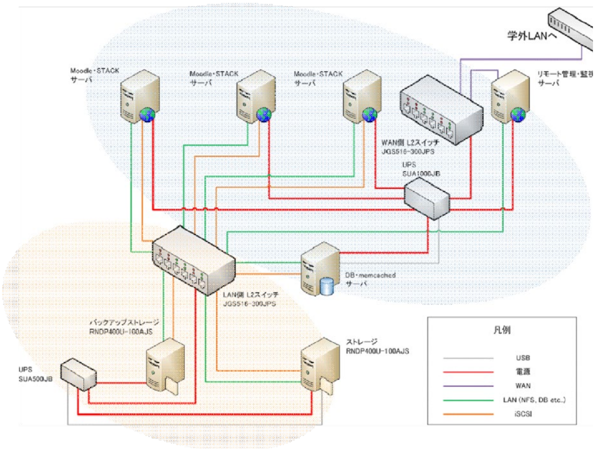


Fig. 8: A conceptual diagram for e-Learning system

4 e-Textbooks

Learning and teaching environment

The e-Learning system utilizes several applications: Moodle, an open-source platform; AMS-LaTeX and MathJax for mathematical descriptions and equation display; STACK and Maxima for online equation assessment in Moodle quizzes; and Geogebra for dynamic mathematics. Figure 9 is the client/server system for e-Learning (Server configuration diagram and actual equipment, Server-side software, In-class and out-of-class learning activities, and client-side web browser).

The learning environment includes a common screen, where students can use their own devices (Bring Your Own Device (BYOD)) under the wireless LAN provided by the university, as well as the desktop computers installed by the university.

The authors, including the teacher and system engineer, provide a server system functioning as aLearning Management System (LMS).

Documents Displaying 2D Mathematical Expressions: Initially, e-Textbooks focused on presenting mathematical expressions in two dimensions. This format allows for the straightforward display of formulas and equations, catering to basic and intermediate levels of mathematics education. Figure 10 is the web-page displaying 2D mathematical expressions (English version transferred by google chrome browser).

Documents with Embedded Dynamic Mathematical Graphs: Advancements have led to e-Textbooks that include embedded dynamic graphs. These interactive elements enable students to visualize and manipulate mathematical concepts, facilitating a deeper understanding of advanced topics through interactive learning.

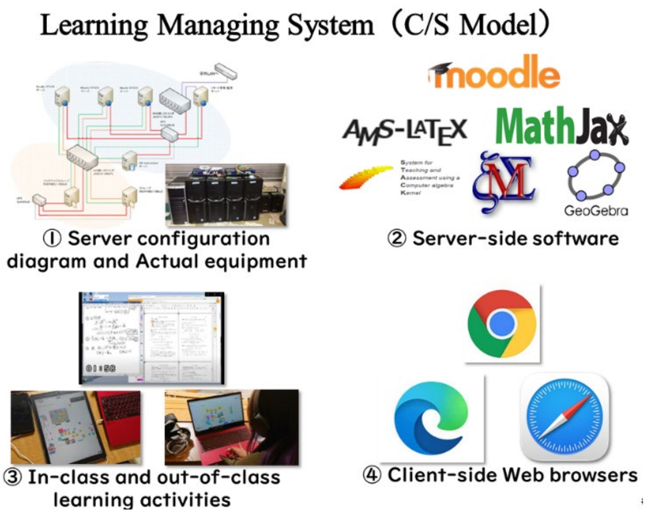


Fig. 9: The concept of client/server system for e-Learning

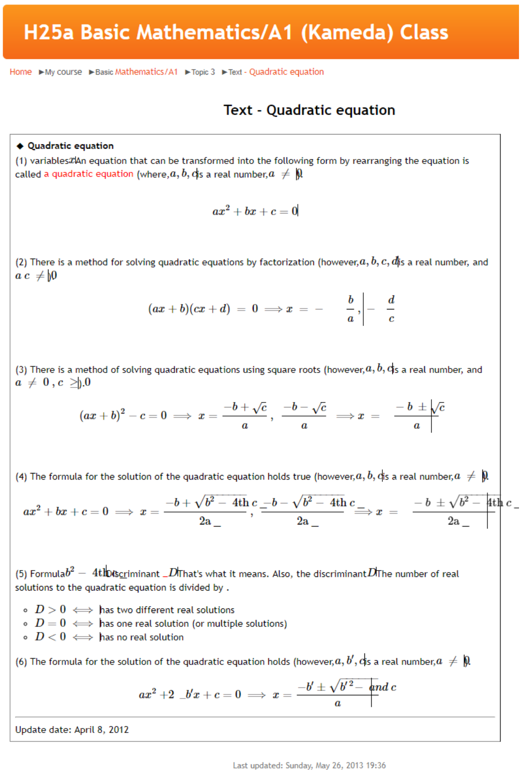


Fig. 10: The web-page displaying 2D mathematical expressions

5 e-Tests

Overview of e-Tests in Mathematics: Covering basic mathematics, differential and integral calculus, and linear algebra.

Fig. 11: Testing at a common place and common time

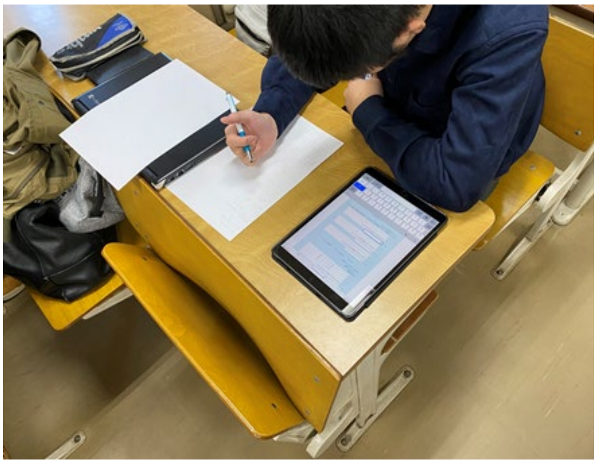


Fig. 12: Testing supported by tablet pc

Classroom-Based e-Tests under direct supervision: These assessments are conducted within a classroom setting, where instructors can directly oversee the testing process to ensure integrity and address any issues in real time (Figures 11–12).

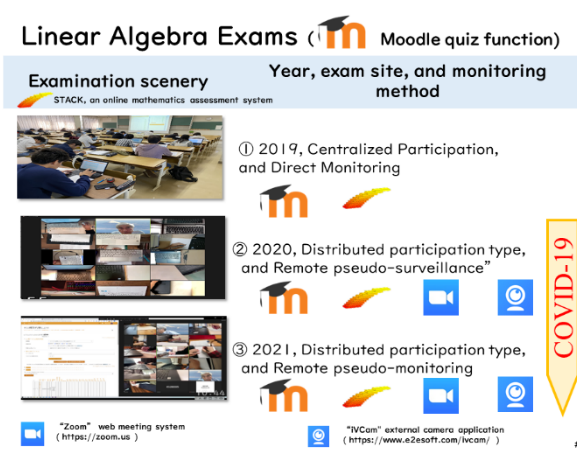


Fig. 13: The transition of the Examination Environment from 2019 to 2022

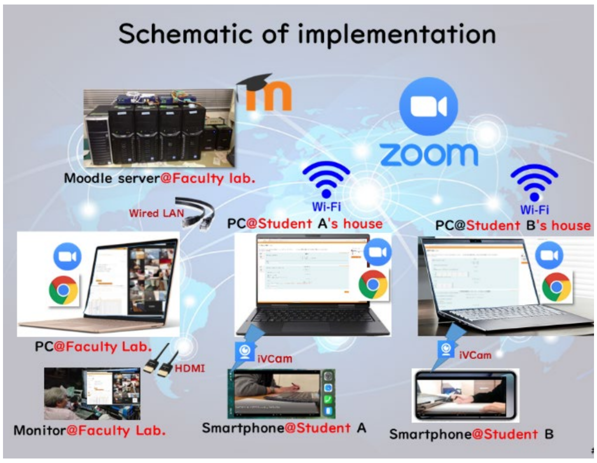


Fig. 14: Remote e-tests were distributed under pseudo-supervision from 2020 to 2021

Distributed e-Tests Under pseudo-supervision: This format involves conducting tests remotely, where supervision is managed through digital means. Although it's not as direct as in-person supervision, various strategies are implemented to maintain the test's integrity ([1], [2], [3]).

Figure 13 is the transition of the examination environment from 2019 to 2022. The exam in 2019 is under direct Supervision, and the exams in 2020 and 2021 is under pseudo-supervision.

Figure 14 is the remote e-tests were distributed under pseudo-supervision from 2020 to 2021. Figure 15 is situation that remote testing was enhanced by utilizing a pseudo-monitoring setup that included a screenshot with two displays: the student's quiz on the left side and a zoom camera gallery on the right side.

Figure 16 is situation that Testing was conducted remotely using a pseudo-monitoring approach, which included a screenshot divided into two sections: the student's initial scoring on the left side and a zoom camera gallery on the right side.

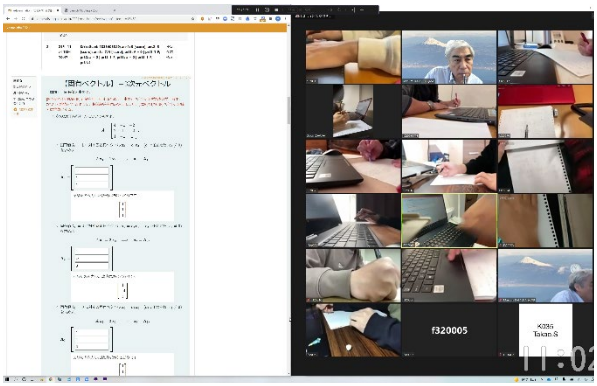


Fig. 15: A view of remoting examination (1)

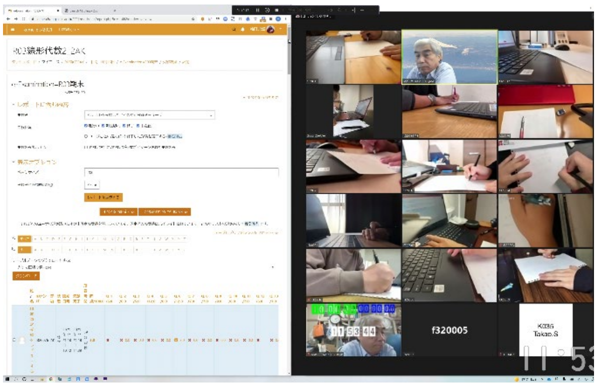


Fig. 16: A view of remoting examination (2)

In 2021, remote testing was enhanced by utilizing a pseudo-monitoring setup that included a screenshot with two displays: the student's quiz on the left side and a zoom camera gallery on the right side. Testing was conducted remotely using a pseudo-monitoring approach, which included a screenshot divided into two sections: the student's initial scoring on the left side and a zoom camera gallery on the right side.

Bibliography

KAMEDA, Masumi (2022). *Practice of non-contact online education for basic mathematics* (in Japanese), Bulletin of the Japan Society for Symbolic and Algebraic , Computation 28 (2), pp. 21–34

KAMEDA, Masumi (2022). *IBT Practice in Linear Algebra* (in Japanese), Bulletin of the Japan Society for Symbolic and Algebraic, Computation 28, pp. 156–159

KAMEDA, Masumi, UDAGAWA, Mitsuru, (2022). *The practice report about the IBT of linear algebra under the examication enviromental of remote pseudo-surveillance and Individually distributed - Using of Zoom, Moodle, etc.* (in Japanese), IPSJSIGTechnicalReport, 2022-CLE-37, pp. 1–6

Focuses: STACK in teaching or exams; Feedback.

Article number: 12

STACK Assessment supported flipped learning model of undergraduate Mathematics in African Universities: a theoretical exploration

Motognon Wastalas d'Assise Dogbalou*; Universita degli Studi di Trieste, Italy

Herine Otieno; EduHubAfric

Abstract

Globally there is growing interest in transforming teaching of mathematics at institutions of higher learning from the traditional lecture method to pedagogical practices that center the learning process on students. This interest is predicated on what has been observed as poor learning results in many of the STEM related fields such as mathematics which have been traditionally taught through the lecture approach. Using a social constructivist lens of teaching of mathematics at university level an explorative theoretical analysis of research on flipped learning and use of STACK for assessment was employed to reflect on the shaping of a STACK assessment supported flipped learning of mathematics at African universities. The paper presents a theoretical position that a STACK assessment supported model of flipped learning of mathematics may help improve the traditional flipped learning model. Specifically, its features such as mastery quizzes at the pre-class session foregrounds increases the opportunities for productive failure in the learning process. Similarly, its features such as the decision trees allow for more responsive teaching informed by real time analysis of key mistakes and misconceptions by the lecturer before the in-class session.

Keywords: pedagogical practices; flipped teaching; productive failure; STACK; technology-assisted teaching; social constructivism; undergraduate mathematics; Africa.

* Corresponding author: wastalasdassise@gmail.com

1. Introduction

This paper is reflective exploration of the opportunities that STACK supported Assessment presents for transformation of teaching of mathematics at African higher institutions of learning from traditional lecture method to more learner centered teaching.

Taking an enactivist reflection framework (Colfes 2012 ; Otieno 2015), the paper seeks to enhance learning, expand potential actions towards pedagogical transformation amongst university mathematics lecturers based on intelligent awareness of the affordances and constraints of STACK present to efforts towards a social-constructivist orientation of teaching of mathematics.

The reflection is premised on emerging lessons from the COVID 19 crisis of 2020 which saw the emergence of STACK supported self-directed learning (Juma et al., 2022) of mathematics by university students in a number of Africa's universities.

Indeed, the restricted interactions occasioned by COVID 19 and difficulties that arose from/with online teaching (Juma, 2023) 'pushed' many of the students to rely on STACK supported formative assessment to facilitate their mastery of several mathematics concepts across different mathematics units.

As such, we surmise that during the COVID 19 crisis of 2020, STACK supported assessment imparted the responsibility of learning on students and temporarily oriented the overall teaching towards a 'quasi' flipped pedagogical approach.

2. Flipped learning

In contrast to traditional model of classroom teaching where students' initial exposure happens in the classroom through a lecture and then later assimilate knowledge through assignments, a flipped model allows students to gain first exposure to new learning material through self-directed reading or watching of videos and after, through discussions and problem solving with teachers and peers assimilate the knowledge (cf. Hao, 2016).

Accordingly, in a flipped classroom, the teacher (lecturer) supports higher order cognitive work by facilitating and guiding students' discussion and problem solving aimed at assimilation and application of the concepts/knowledge.

Indeed, as one of the active learning pedagogies, flipped learning is increasingly being adopted by mathematics department in several universities in the West (look for papers on mathematics education on flipped learning from west) and increasingly in the East (look for papers on mathematics education on flipped learning from East)) to shift learning of mathematics from passive to active learning thereby promoting ownership of learning and fostering of deeper, more visible, reflexive and collaborative learning amongst students.

Instructively, despite the aforementioned advantages of flipped pedagogy, there is very little knowledge and practise of the same in the African mathematics education context.

Further, a majority of the extant research focus on flipped learning programs for mathematics at higher institution of learning where recorded videos is the main medium through which students engage individually with the new concepts before interacting with peers and their lecturers (Baker, 2000). Notably, a video model format of flipped learning may not fit the African context given the technology related challenges for both learners and lecturers.

Drawing on evidence from related research on self regulated learning of mathematics education by secondary school students in Kenya students (Otieno & Povey 2022) which espouses quality of mathematics questions and worked out examples as central to their self directed learning of mathematics concepts , we propose to theoretically explore the opportunity that STACK assessment provides for shaping a contextually responsive flipped learning of mathematics in African higher institutions of learning.

Accordingly, this paper seeks to contribute to filling the paucity of research on flipped learning approach in the African context and the knowledge gap on non-video supported flipped learning of the mathematics at higher institutions of learning.

The theoretical and reflective exploration is guided by the following research questions :

- How may features of flipped learning map into features of STACK Assessment model of flipped learning ?
- What are the possible solutions to the challenges that African Universities may face in implementing STACK supported flipped pedagogy for mathematics?

3. Methodology

This a theoretical exploratory study : given the paucity of research on an individualised assessment model of flipped learning for mathematics in African Institutions of higher learning, The authors draw from the enactive reflection from their experience of supporting use of STACK assessment in African Universities to analyse and make connections from empirical studies on flipped learning to examine, define and advance a STACK Assessment supported model for flipped learning of mathematics in the African universities context.

4. Findings

Mapping of traditional flipped learning features to a STACK Assessment mode of flipped learning

	Traditional Flipped learning features	Equivalent feature on a STACK Assessment model of flipped learning of mathematics African Universities
Content	Lecturer provides accessible and relevant content mostly in form of videos to students (Milman, 2013). This may be accompanied by text readings and quizzes aimed at aiding students conceptual understanding(Baker, 2000).	Class notes or key references for target unit together with topic based automated assessment (Mora, 2020) and worked solutions of questions as learning-for the purposes of supporting students' self-directed
Engagement	Students listen to videos, reflect on their own learning engage in peer discussion and teach each other . (Chowdhury et al., 2019 ; Reddan et al. (2016).	Students review class notes, related resources and worked solutions of examples provided through STACK.. Students engage in peer teaching partly faciltated through STACK mastery quizess Engagement is stregthened by low stakes feature of STACK mastery

		questions since it allows for multiple attempts without grading or ranking (Wieman, 2007; Chowdhury et al., 2019)
Student-Teacher Interaction	Mostly happens back in the classroom where the teacher facilitates discussion amongst students and with students on areas of difficulty experienced by the students individually or collectively (Bergman and Sams, 2012)	STACK can promote a post initial exposure student-teacher interaction through a number of its features : Use of official discussion forums and chat features or complimentary social media platforms for communication between students and teachers : students may raise questions seeking teachers support or just flag out areas of difficulty that they want teacher to focus on during the lesson. Teacher may tailor in class session to facilitate discussion amongst the students and with students to address some of the gaps in understanding unearthed in following the provided examples or trying out the mastery quizzes. During in person session address mistakes and misconceptions identify from the analysis of data from students attempts at the mastery quizzes .
Assessment	*Assessment at the difference stages:* Self -assessment which may take the form of quizzes at different stages of learning including for the purposes of content mastery and higher order thinking skills such as problem solving and critical thinking (McLaughlin et al. 2014).	*Formative assessment:* • **Mastery and Test Quizzes:** topical quizzes that can be repeatedly attempted by students until they gain a specified level of mastery. This iterative process helps ensure that students have thoroughly grasped the material before moving on to more advanced topics. By providing instant feedback, STACK helps students identify and correct their mistakes in real-time (Sangwin et al., 2019). • **Unit Assessment:** STACK supports formative assessment by allowing instructors to create quizzes that give immediate feedback on students'

		understanding. This type of assessment helps inform both students and teachers about areas that need further attention and provides ongoing insights into student progress. The ability to tailor questions and feedback to individual student needs enhances personalised learning (Juma et al. 2022). *Summative Assessment:* • STACK is used for summative assessment by creating detailed tests that assess students' overall knowledge at the end of a learning period. Automated grading streamlines the assessment process, making it efficient for large classes. • **Exams**: STACK facilitates the creation and administration of end semester exams that rigorously assess student understanding under controlled conditions. The system supports a wide variety of question types, from multiple-choice to complex mathematical problems, ensuring a robust evaluation of student competence. The secure environment provided by STACK ensures the integrity and fairness of the examination process.
Feedback	The pre-class learning activity in flipped learning is often followed by formative assessment and feedback activities. The drawing on the pre class quiz to identify students' possible misconceptions helped in reinforcing in class teaching to support deeper learning amongst the students. That said, often, teachers are not able to track and synthesise the misconcpetions	As alluded to in the assessment section, STACK provides various avenues for providing feedback on students learning right from pre-class learning session and after class learning session. Indeed, STACK features such as provision of real time feedback based on errors made by students can support learning when using a flipped learning pedagogy both at pre and in class stage.

	before the in-class session (Kapur et al., 2022).	That the feedback is delivered through Potential Response Trees (PRTs), allowing for tailored guidance, and the inclusion of worked solutions further enhances understanding by offering opportunity for the lecturer to provide step-by-step explanations. (Santiago, 2023)

The diagram below maps STACK assessment features as currently used by university students in some of African universities to flipped learning component.

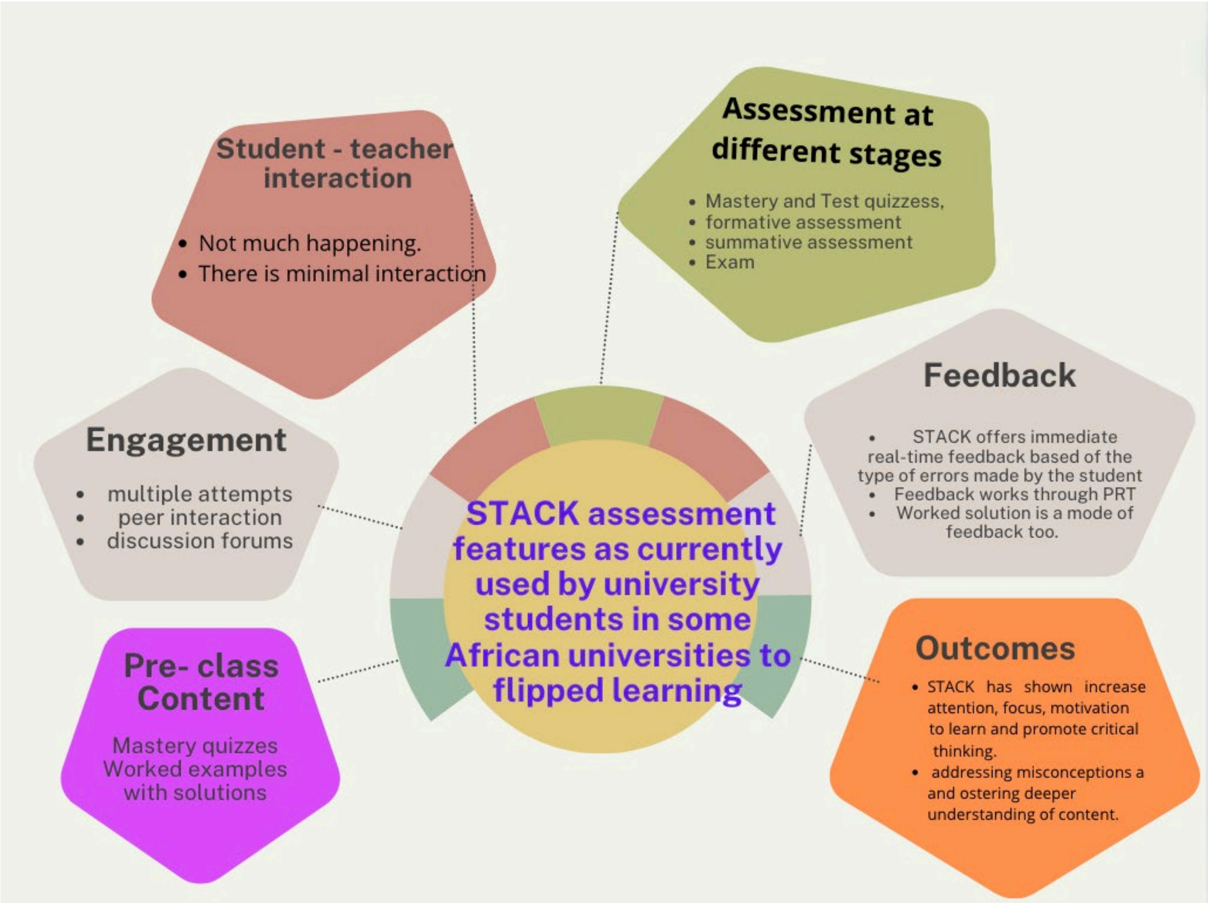


Fig. 1: Maps STACK assessment features as currently used by university students in some of African universities to flipped learning component

Possible solutions to the challenges that African Universities may face in implementing STACK supported flipped pedagogy for mathematics?

To effectively implement flipped pedagogy with STACK, teachers must address several key challenges. These include mastering the necessary technological skills, managing the time required for material preparation, ensuring students are ready for self-directed learning, maintaining student engagement with pre-class materials, and adapting in-class activities to build on prior learning (Ustinova et al., 2020).

Technological Proficiency: Teachers may face difficulties in mastering the technical aspects of STACK. Providing comprehensive training sessions and ongoing technical support can help teachers become proficient in using STACK effectively.

Time Management: Developing and organising flipped classroom materials requires significant time investment. Collaborating with colleagues to share resources and using pre-existing STACK question banks can reduce the workload. Additionally, setting aside dedicated time for planning can help manage this challenge.

Student Readiness: Students might not be accustomed to self-directed learning and could struggle with the initial transition. Implementing gradual changes to the teaching approach and providing clear instructions and guidance can ease students into the flipped model. Encouraging a growth mindset and offering support for time management and study skills can also be beneficial.

Engagement and Participation: Ensuring all students are actively engaging with the pre-class materials can be challenging. Incorporating interactive and varied content in the pre-class materials, such as worked examples, class notes, quizzes, and discussions, can increase student engagement. Regular checks and formative assessments can also help monitor and encourage participation.

Classroom Dynamics: Adapting in-class activities to ensure they build on the pre-class work and promote active learning can be difficult. Designing interactive and collaborative in-class activities that require the application of pre-learned concepts can enhance learning. Using group work, problem-solving tasks, and discussions can foster a more dynamic and engaging classroom environment.

5. Conclusion

Evidence from research on flipped learning suggest that there is no one method of ‘flipping’ and that there is still more to understand about the phenomenon not just in Africa but across different learning contexts globally (cf. Garrison et al. 2007).

Accordingly, our proposal for a STACK assessment supported flipped learning of mathematics is important not in seeking the above gap but also gaps that have been identified in traditional flipped learning approach: there has been concerns that most flipped learning fail to achieve its key goal of exposing students new concepts, helping identify what the students know and do not know and drawing on this knowledge of understandings and misunderstandings to inform in class teaching (Hew and Lo 2018).

Indeed, by proposing a model of flipped learning that foregrounds engaging with questions and worked examples as key content of pre-class session, we will reinforce the opportunity for learning by drawing on the affordances of productive failure (Schneider et al., 2013; Hu et al., 2019) to facilitating learning. Further, tapping into the features of STACK assessment that allows for real time feedback and analysis of students errors during pre-class individual and peer learning sessions there is enhanced opportunities for teachers to tap into the students’ misunderstandings and conceptions to improve teaching.

As such a STACK supported model of flipped teaching of mathematics may not only help circumvent inherent technological challenges that may come with using videos for pre class sessions, it may actually lead to an improved and more effective flipped learning. In essence, a STACK assessment supported flipped learning may resonate closely to evidence from related research that has called for a rethinking of flipped learning to a fail, flip, fix and feed model (Kapur et al. 2023).

Future research is needed to investigate the validity, reliability and impact of the STACK supported flipped learning to teaching and learning outcomes of mathematics in African universities.

Bibliography

Arum, R., & Roska, J. (2011). *Academically adrift: Limited learning on college campuses*. University of Chicago Press.

Bach, S. (2020). Using multiple choice questions in STACK – reasons and examples. https://doi.org/10.5281/ZENODO.3945809

Baker, J. W. (2000). The classroom flipped: Using web course management tools to become the guide on the side. Paper presented at the 11th International Conference on College Teaching and Learning, Jacksonville, FL.

Garrison, D. R., & Kanuka, H. (2004). Blended learning: uncovering its transformative potential in higher education. *Internet and Higher Education, 7*(2), 95–105.

Hao, Y. (2016). Exploring undergraduates' perspectives and flipped learning readiness in their flipped classrooms. *Computers in Human Behavior, 59*, 82–92. https://doi.org/10.1016/j.chb.2016.01.032

Hew, K. F., & Lo, C. K. (2018). Flipped classroom improves student learning in health professions education: A meta-analysis. *BMC Medical Education, 18*, 38. https://doi.org/10.1186/s12909-018-1144-z

Hu, X., Zhang, H., Song, Y., Wu, C., Yang, Q., Shi, Z., et al. (2019). Implementation of flipped classroom combined with problem-based learning: An approach to promote learning about hyperthyroidism in the endocrinology internship. *BMC Medical Education, 19*, 290. https://doi.org/10.1186/s12909-019-1714-8

Juma, Z. O. (2023). Analysis of the impact of the stack system on learner engagement, performance, and perception on its use in mathematics at Maseno University, Kenya (Master thesis). Maseno University, Kenya.

Juma, Z. O., Ayere, M. A., Oyengo, M. O., & Osang, G. (2022). Evaluating engagement and learning based on a student categorization using STACK, exam data, key informant interviews, and focus group discussions. *International Journal of Emerging Technologies in Learning (iJET), 17*(23), Article 23.

Juma, Z. O., Okoth, I. O., & Oyengo, M. O. (2022). Delivery of integral calculus at Maseno University: Is STACK really playing an integral part? *International Journal of Emerging Technologies in Learning (iJET), 17*(23), 64–68. https://doi.org/10.3991/ijet.v17i23.36621

Kemmis, S. (2021). Assessment OF/FOR/AS Learning - It's About Learning. Retrieved from https://www.harapnuik.org/?page_id=8900

Kim, M. K., Kim, S. M., Khera, O., & Getman, J. (2014). The experience of three flipped classrooms in an urban university: An exploration of design principles. *The Internet and Higher Education, 22*, 37-50.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

Love, B., Hodge, A., Grandgenett, N., & Swift, A. W. (2014). Student learning and perceptions in a flipped linear algebra course. *International Journal of Mathematical Education in Science and Technology, 45*(3), 317–324.

McLaughlin, J. E., Roth, M. T., Glatt, D. M., Gharkholonarehe, N., Davidson, C. A., Griffin, L. M., et al. (2014). The flipped classroom: A course redesign to foster learning and engagement in a health professions school. *Academic Medicine, 89*(2), 236–243.

Milman, N. B. (2013). The flipped classroom strategy: What is it and how can it best be used? *Distance Learning, 9*(3), 85–87.

Otieno, H. (2015). Kenya secondary school students' intelligence beliefs: A case study in mathematics. *Proceedings of the British Society for Research into Learning Mathematics, 35*(2), 35-2.

Otieno, H., & Povey, H. (2023). Mathematics textbooks and self-regulated learning: responses from students in three Kenyan secondary schools. *Research in Mathematics Education, 25*(3), 342-358.

Oyengo, M. O., Parsons, D., Stern, D., & Sangwin, C. (2021, June 9). Providing student feedback through electronic assessment for linear algebra at Maseno University, Kenya. *International Meeting of the STACK Community 2021*. Zenodo. https://zenodo.org/records/5035969

Sangwin, C. J., Hancox, J. E., Heeffer, A., & Trgalova, J. (2019). Adapting to students' solutions with STACK. *International Journal of Mathematical Education in Science and Technology, 50*(6), 823-836.

Schneider, B., Wallace, J., Blikstein, P., & Pea, R. (2013). Preparing for future learning with a tangible user interface: The case of neuroscience. *IEEE Transactions on Learning Technologies, 6*(2), 117–129. https://doi.org/10.1109/TLT.2013.15

Tang, T., Abuhmaid, A. M., Olaimat, M., Oudat, D. M., Aldhaeebi, M., & Bamanger, E. (2020). Efficiency of flipped classrooms with online-based teaching under COVID-19. *Interactive Learning Environments*, 1-12.

Ustinova, I., Tomilenko, V., Imas, O., Beliauskene, E., & Yanuschik, O. (2020). STACK assessment in mathematics classroom: Advantages and disadvantages. *International Conference on Interactive Collaborative Learning*, 174-182. https://doi.org/10.1007/978-3-030-68198-2_15

Zhao, L., He, W., & Su, Y.-S. (2021). Innovative pedagogy and design-based research on flipped learning in higher education. *Frontiers in Psychology, 12*, 577002. https://doi.org/10.3389/fpsyg.2021.577002

Focuses: use of (interactive) JSXGraphs; generation of feedback; randomized questions; assessment and self-study.

Article number: 13

STACK for PDEs

George Ionita, Laura Kobel-Keller*; ETH Zurich

Florian Spicher; Universität Bern

Abstract

In this joint project we extend the range of STACK exercises to Partial Differential Equations (PDEs), setting new standards in this domain. Beyond the comprehensive integration of PDE exercises into STACK, a pioneering achievement in itself, our technical implementation introduces innovative features, such as JSXGraph.

The diverse exercises provide students with a feedback-rich training environment to practice techniques to solve PDEs - common in natural and engineering sciences — using varying data until mastery is achieved. In addition, these exercises also offer opportunities to explore novel features, such as the mathematical description of the hydrogen atom.

These tasks earned high appreciation from the students and engaged them more actively in comparison to traditional pen-and-paper exercises. Notably, the individual feedback in the response trees proved crucial for students in their learning journey, particularly in large classes where these interactive exercises serve as a partial substitute for teaching assistants.

* Corresponding author: laura.kobel-keller@math.ethz.ch

1. Introduction/Background

In recent years, the student numbers grew rapidly while the supervision/mentoring capacities (lecturers and teaching assistants) remained rather constant. This led to bigger and bigger classes where individual mentoring and guiding is no longer possible.

One potential solution to this challenge is the use of interactive exercises with individual feedback. On Moodle, the platform most commonly used for lecture websites, STACK provides such a feature. STACK – which is also available in ILIAS and other similar platforms – provides a feedback rich, interactive and randomized exercise environment supported by the robust computer algebra system MAXIMA. This powerful tool aids students in their learning process while providing lecturers with valuable insights into the areas where students face the most difficulty.

STACK has already been successfully used at various universities, including ETH Zürich, starting from the mathematics lectures in the first year (Basisjahr) and by now ranging also to other, non-mathematics courses.

In our project, we extend these concepts to second year mathematics contents, in particular Partial Differential Equations (PDEs). Setting thus not only new standards in the mathematical education, but also in the community of STACK users.

The topic of Partial Differential Equations (PDEs) is fundamental for all students in natural and engineering sciences being one of the most common ways of stating a mathematical model. Mastering and solving such equations requires routine and extensive practice – often more than what traditional exercise sheets can offer. One key advantage of STACK exercises over classical exercises is that they allow that such calculation routines can be easily trained by randomized exercises, which is one of the core features of STACK.

2. Covered mathematical topics and mission

For all the following topics we wrote at least one representative exercise with various randomized variants, elaborate response trees and detailed model solutions. In addition, to enhance the learning effect, we gave as little hints about the technique that should be used as possible, e.g. we tried not to indicate a corresponding keyword in the question title or in the introductory text.

We adopted the setting of functions in two variables for the study of PDEs because "engineering techniques" leading to explicit solutions are most clearly presented in this context. Additionally, students are most familiar with this setting from their previous lectures. More precisely, our setting was one spatial and one time variable.

The primary aim of this initial(first) set of questions was to deepen and consolidate the students' calculation routines. It also provided ample feedback and support during their learning journey, even in the absence of in-person tutoring.

This set covers the following mathematical topics:

- Separation of variables (for wave and heat equation)
- Fourier transform
- Laplace transform
- Duhamel principle
- D'Alembert formula

- Propagation and separation of waves
- Travelling waves
- Method of characteristics
- Non-homogeneous problems with several possible solution approaches

In the second set of STACK questions, we shifted the focus and design to better support self-study and exploration of new material. To accomplish this, we included more introductory text, provided clear and precise instructions on the calculations to perform, and offered extensive references to helpful resources.

These questions focus on modeling the hydrogen atom. Starting with the Schrödinger equation and Coulomb's law, we guided the students through reducing the problem to a single variable (the radial one in terms of the actual problem). This process ultimately led to the Laguerre equation, with the solutions being the corresponding Laguerre polynomials.

3. Special features

Based on the literature and extensive discussions with colleagues and experts in the field, our collection of STACK exercises on PDEs appears to be unique. This uniqueness stems from our utilization of the latest features of STACK, comprehensive coverage of the material, and the high quality of the exercises.

4. Some examples

In the following, we will delve into the specific characteristics of some of our PDE exercises. These exercises are part of a sequence of three on traveling waves, each building upon knowledge from the previous exercise. Together, they form a cohesive unit in terms of both form and content, effectively integrating several key features.

First example: Most elaborate response tree, in particular in the given feedbacks

The first exercise aims to introduce the topic to students in a gradual manner, by recalling the relevant theory and formulas, as well as the notations. The explicit text of this exercise can be seen in Figure 1.

Given simple boundary functions f and g, where f is compactly supported with a random value a (a random variable) and support $[-h, h]$ (again h being a random variable) and $g = 0$ is fixed, the students are asked to compute the exact expressions of the corresponding traveling waves F and G. And the information that the final solution of the given PDE problem can be written as a superposition of the two traveling waves F and G (one traveling to the right, one traveling to the left) is recalled in the text of the exercise.

STACK question dashboard

The aim of this exercise is to analyze and understand the motion of an infinite string.

This motion is modeled by the wave equation $u_{tt} = c^2 u_{xx}$ - where $x \in \mathbb{R}$ and $t \geq 0$ - paired with the following initial conditions

$$u(x,0) = f(x) = \begin{cases} 9, & \text{if } |x| \leq 10 \\ 0, & \text{otherwise} \end{cases}$$

and

$u_t(x,0) = g(x) = 0$, for all $x \in \mathbb{R}$ and all $t > 0$.

Recall that from d'Alembert's formula, the solution of this problem is of the form

$$u(x,t) = F(x-ct) + G(x+ct),$$

where F and G are a right- and a left-moving wave travelling at speed c, respectively, and given by

$$F(x) = \frac{1}{2}f(x) - \frac{1}{2c}\int_0^x g(s)\ \mathrm{d}s \text{ and } G(x) = \frac{1}{2}f(x) + \frac{1}{2c}\int_0^x g(s)\ \mathrm{d}s.$$

Compute F and G.

$F(x) =$ ____ , if ____

$F(x) =$ ____ , otherwise.

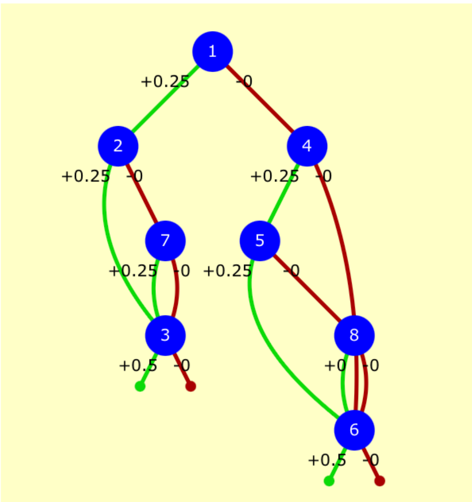
and

$G(x) =$ ____ , if ____

$G(x) =$ ____ , otherwise.

Figure 1: Exercise text

In this exercise, the specialty is the most elaborate response trees, PRT (see also Figure 2):



1. Checks if the first value is aa/2	`ATAlgEquiv(ans1,ta1)`	prt1-1-T prt1-1-F
2. Checks the range associated to aa/2	`ATAlgEquiv(ans2,ta2)`	prt1-2-T prt1-2-F
3. Checks the second value	`ATAlgEquiv(ans3,ta3)`	prt1-3-T prt1-3-F
4. Checks if the first value is 0	`ATAlgEquiv(ans1,ta3)`	prt1-4-T prt1-4-F
5. Checks the range associated to the value	`ATAlgEquiv(ans2,ta2_reversed)`	prt1-5-T prt1-5-F
6. Checks if the second value is aa/2	`ATAlgEquiv(ans3,ta1)`	prt1-6-T prt1-6-F
7. Alternative form of range	`ATAlgEquiv(ans2,ta2_alternative)`	prt1-7-T prt1-7-F
8. Alternative form of range	`ATAlgEquiv(ans2,ta2_alternative_reversed)`	prt1-8-T prt1-8-F

Figure 2: Response tree (for the first three answers) of the above displayed question

Since $F = G = \frac{1}{2}f$, both functions, F and G behave similarly. Specifically, it is checked whether the main values of the waves are correct (node 1) and whether the supports are the same as f (nodes 2,7) in both the absolute value notation ($|x| \leq h$) and with two inequalities ($-h \leq x \leq h$). If the main values are incorrect, it is verified whether the function is written as zero in a certain range and $\frac{1}{2}a$ in the complement (nodes 3,4,5,6,8).

Note that in this PRT is designed in such a way that either possible way the students enter the functions $F(x)$ and $G(x)$ - provided the given information is correct - they receive full credit. More precisely, if a student first indicates where one of these function takes the value zero and then in a second step indicates where this same function takes the non-zero value, full credit is given.

This is also reflected in the specific feedback and partial credits the students are given. They receive individual, detailed feedback and partial credit on the values of the corresponding function and (independently) on the respective intervals.

Second example: Use a JSXGraphs, one in a static fashion, one in an interactive fashion

The second exercise, which focuses on the geometry of travelings waves, aims to train students in reading characteristics and deducing information about the solution u of the given PDE problem. Given a proved, static graph, the students must place three points among six distinct regions. By deducing the values of the traveling waves F and G in those regions, they are led to the desired values for u.

Importantly, the numerous possible points satisfy some constraints, even though they are chosen arbitrarily from a list of possibilities, which depends on the random variables a, h c. Specifically, the points are selected as follows:

1. Each point belongs to a different region;
2. At least one of them is not in regions IV-VI, ensuring that u is nonzero at that point;
3. $x \neq 0$ and $t > 0$ are taken such that the point (x, t) does not lie on the characteristic lines.

The special feature of this exercise is the (static) JSXGraph that is given to the students together with the exercise text itself in order to them recalling/learning the graphic picture of the situation of traveling waves.

Here again, the reader can find the whole exercise text in Figure 3.

STACK question dashboard

The aim of this exercise is again to analyze and understand the motion of an infinite string, but this time by looking more into geometric properties of the domain on which the problem should be solved.

Recall that we consider the motion of an infinite string modeled by the wave equation

$$u_{tt} = c^2 u_{xx},$$

where $x \in \mathbb{R}$ and $t \geq 0$, paired with the following initial conditions

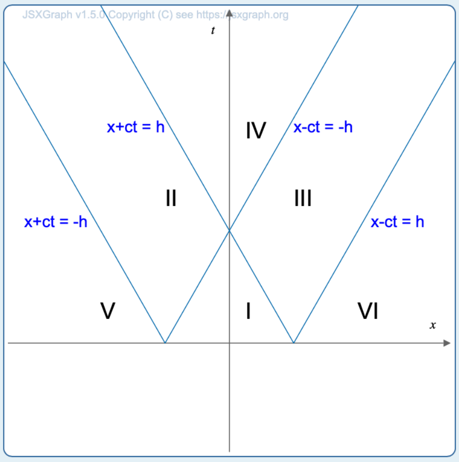
$$u(x,0) = f(x) = \begin{cases} 1, & \text{if } |x| \leq 5 \\ 0, & \text{otherwise} \end{cases}$$

and

$$u_t(x,0) = g(x) = 0, \text{ for all } x \in \mathbb{R} \text{ and all } t > 0.$$

Note that outside the interval $[-h,h] = [-5,5]$ the initial condition vanishes and the maximal value of the initial condition is $a = 1$.

According to the notion of travelling waves (see also the domain of influence), $\mathbb{R} \times [0,\infty)$, which is the domain of our solution $u(x,t)$, can be divided into the following regions displayed below.



If $c = 8$, $a = 1$ and $h = 5$, i.e., if $u_{tt} = 64\,u_{xx}$ with boundary conditions $u(x,0) = f(x) = \begin{cases} 1, & \text{if } |x| \leq 5, \\ 0, & \text{otherwise,} \end{cases}$ and $u_t(x,0) = g(x) = 0$, then determine for each of the following points (x,t) in which region it lies and the corresponding value of $u(x,t)$:

The point $(x,t) = \left(-\frac{5}{4}, \frac{5}{4}\right)$ lies in region Ex: VII , and $u\left(-\frac{5}{4}, \frac{5}{4}\right) =$

The point $(x,t) = \left(-5, \frac{5}{32}\right)$ lies in region Ex: VIII , and $u\left(-5, \frac{5}{32}\right) =$

The point $(x,t) = \left(\frac{5}{3}, \frac{5}{32}\right)$ lies in region Ex: IX , and $u\left(\frac{5}{3}, \frac{5}{32}\right) =$

Figure 3: Use of a static JSXGraphs element

The corresponding PRTs simply check whether each point belongs to the correct region and leads to the appropriate value for the solution u.

Finally, in the third exercise of this sequence, we want to assess the students' understanding of the physics of the problem with a final part based on the motion of the traveling waves. The students are first asked to identify/calculate the time of separation of the waves t_s using the formula $t = \frac{d}{v} = \frac{2h}{2c}$. The previous parts should help them finding the distance $d = 2h$ at which the waves separate, and they are also expected to determine the speed of separation $v = 2c$ as F and G move apart.

Next, the students are given a new time T and four interactive graphs that may describe the motion of the solution u. As T is chosen among four scenarios (slightly/considerably smaller/bigger than t_s), the students need to use a cursor on the graphs, showing the evolution of the solution from $t = 0$ to $t = T$, and to select which one represents the correct situation, i.e., which graph shows the evolution of the solution u up to $t = T$.

Also here, the full text of this exercise is given here, namely in Figures 4 and 5.

The aim of this exercise is again to analyze and understand the motion of an infinite string, but this time by looking more into geometric properties of the domain on which the problem should be solved. STACK question dashboard

Recall that we look at the motion of an infinite string modeled by the wave equation

$$u_{tt} = c^2 u_{xx},$$

where $x \in \mathbb{R}$ and $t \geq 0$, paired with the following initial conditions:

$$u(x,0) = f(x) = \begin{cases} 3, & \text{if } |x| \leq 3 \\ 0, & \text{otherwise} \end{cases}$$

and

$$u_t(x,0) = g(x) = 0, \text{ for all } x \in \mathbb{R} \text{ and all } t > 0.$$

Note that outside the interval $[-h, h] = [-3, 3]$ the initial condition vanishes and the maximal value of the initial condition is $a = 3$.

According to the notion of travelling waves (see also the domain of influence), $\mathbb{R} \times [0, \infty)$, which is the domain of our solution $u(x,t)$, can be divided into different regions (see also picture in the exercise "Infinite string - Travelling waves (part 2)").

Now, say we have $c = 7$, $a = 3$ and $h = 3$, i.e., if

$$u_{tt} = 49\, u_{xx}$$

with boundary conditions

$$u(x,0) = f(x) = \begin{cases} 3, & \text{if } |x| \leq 3, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$u_t(x,0) = g(x) = 0.$$

Question 1. Compute the time of separation t_s of the two pulses F and G.

Hint: This time t_s can also be interpreted as the first time such that $u(0, t_s)$ vanishes.

t_s =

Figure 4: First part of the exercise where interactive JSXGraphs elements are used

Question 2. Based on your previous result, find the graph corresponding to the solution $u(x, t = T)$, with $T = \frac{69}{140}$ s.

Hint: You should choose one of the pictures below where they are numbered from 1 to 4 starting on top. You might also use the slider in order to look at different times.

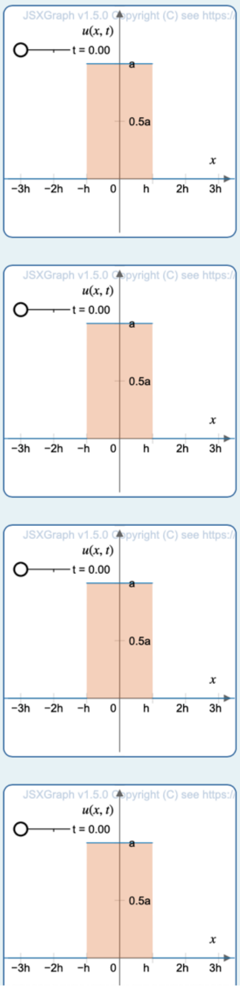


Figure 5: Second part of the above mentioned question with the interactive JSXGraphs elements

An initial standard programming of this exercise led to a randomized problem in the variables with fixed positions of the graphs, thus always describing the same situation, meaning the correct solution would not change its location. This caused our question to quickly become uninformative for the students. To solve this issue, we included randomization for the position of the graphical solution too. Consequently, we had to challengingly include these two randomization processes in the PRT verifications and feedback. This is done by extracting the time represented in the graph passed as an answer and comparing it to t_s.

Acknowledgment

The authors thank the Department of Mathematics of ETH Zürich for having enabled this project and the generous financial support.

Bibliography

Sangwin, Ch. (2007). *STACK: making many fine judgements rapidly*. CAME. https://api.semanticscholar.org/CorpusID:62695279

Sangwin, Ch. (2015). *Computer Aided Assessment of Mathematics Using STACK*. Selected Regular Lectures from the 12th International Congress on Mathematical Education, pp. 695-714. https://doi.org/10.1007/978-3-319-17187-6

Glasmachers, E., Kallweit, M. (2019). *Adaptive Selbstlernaufgaben mit STACK*, Contributions to the 1st International STACK conference 2018 https://doi.org/10.5281/zenodo.2562732

Knaut, J., Altieri, M., Bach, S., Strobl, I., & Dechant, K. (2022*). A Theory-Based Approach of Feedback in STACK-Based Moodle Quizzes Taking into Account Self-Regulation and Different Proficiency of Learners.* International Journal of Emerging Technologies in Learning (iJET), 17(23), pp. 38–55. https://doi.org/10.3991/ijet.v17i23.36425

Information on the authors

George Ionita

George Ionita has been working with STACK for the past 5 years. Currently, he is part of the E-Learning Group as Digital Teaching Specialist at the Department of Mathematics at ETH Zürich.

Laura Kobel-Keller

Laura Kobel-Keller has been a Senior Scientist ETH Zürich since 2016.

Apart from her teaching activities, she does research in the area of partial differential equation and in the domain of didactics of mathematic at university level.

Florian Spicher

After his Master's degree in Applied Mathematics at ETH Zürich in 2023, he engaged in teaching and educational development activities at ETH Zürich and is now pursuing his PhD at the University of Bern.

Focuses: Feedback; STACK in teaching or exams.

Article number: 14

STACK in physics across the curriculum

Abigail Pied*, Alastair Palmer, Eric Yao

University of Glasgow

Abstract

Diversifying and distributing assessments is useful in promoting student engagement and enhancing inclusiveness [OP22]. Computer-based automated marking online assessment provides a key ingredient for doing this in large class settings [PMH22]. It can promote engagement by increasing time spent by students on learning tasks, providing them with feedback and valuable practice opportunities. Because it is largely free of the need for teacher intervention at specific time, it allows for asynchronous learning to take place. However, typically, teachers are concerned about the scope and range of the questions that can be set for these types of assessments and students are worried that their work is not fairly assessed due to limitations of the input system used. Although designed for testing mathematics, STACK is well suited for posing physics questions. It has been adopted by physics departments at various universities, frequently for junior levels where numerical calculations and simple derivations can be tested. Here we present our adaptation of exam and tutorial questions across the levels of the physics degree curriculum while addressing common perceptions of shortcomings of computer-based automated marking assessments. In this work, STACK provides the foundation for the assessment while being supplemented by other Moodle question types, Moodle being the VLE adopted in our institution. To address possible teacher concerns, we start from "standard" question set which tests a wide range of intended learning outcomes and cognitive skill levels. In deploying these questions, we focus on providing fairness and consistency in marking as well as timely and personalised feedback. By encouraging student to simultaneously upload their workings and analysing student responses, accurate assessment of the student attempts can be made. Deploying these types of assessments both as formative and summative assessments in various classes shows good correlation with student performance in "traditional" assessment types. This means that STACK, when used in conjunction with other tools, can be an extremely useful tool in learning and teaching in physics and other fields.

Keywords: physics, e-assessment, feedback, STACK, Moodle

* Corresponding author: abigail.pied@glasgow.ac.uk

1. Introduction

The use of computer-based assessment has developed a lot in the past decades, particularly in recent years due to the pandemic, as online assessment became the norm during that period. With its use becoming more widespread and normalised, it is worth investigating how computer-based assessment compares to traditional handwritten assessment; in terms of its effectiveness at assisting the students' learning, its ability to deliver useful feedback, and the nature and complexity of the questions the students are given. In this paper, we first present how we have approached adapting pre-existing physics questions into STACK questions, then we explain the process used to deliver highly personalised feedback to the students, and finally we have a preliminary look at the effectiveness of physics computer-based homework at helping students prepare for a class test.

2. Writing STACK questions for physics

Adapting preexisting physics problems into STACK questions

STACK was chosen to be used in the homework and assessment of the students for its compatibility with STEM questions, and its ability to create multi-part questions. One of the goals of our work with STACK is to assess its suitability as a tool to pose physics questions, and its ability to assess the same skills as a traditional physics problem. Consequently, we wanted our STACK questions to be based on existing traditional physics problems, rather than create questions designed specifically for STACK and its functionalities, to make comparison possible.

When writing physics questions in STACK, the aim was to conserve as much of the question as possible, with minimal modifications. This means that not all questions were possible to adapt into a Moodle quiz using STACK exclusively. Indeed, students will often be asked to give a definition, describe a process, or explain the significance of an experiment, for example. This cannot be included inside a STACK question without significantly altering the complexity of the problem, for example, by substituting it with a fill in the blank paragraph with drop-down menus. Instead, it was decided to split such questions into an essay type question, where students can type their answer which will be manually assessed after submission, and a STACK question for the rest of the question. Similarly, students are sometimes asked to sketch a diagram or a curve, and while some of these can be included into a STACK question, using such tools as Meclib [KS21] for free-body diagrams and JSXGraph, it is not always possible, or sometimes too time-consuming to code for a single question. Then, such problems can also be adapted using the essay question type to allow students to upload a picture of their answer. This approach to adapting physics problems does mean that if a question makes use of an essay type question it will prevent the use of immediate or adaptive feedback.

Overall, we have found that most parts of a physics problem can be adapted into a STACK question without much, or any need for modification. Indeed, at the simplest level, we have plain numerical applications which can be tested checking for number of significant figures, percent error from the expected result and units in STACK. Short or simple derivations can usually be tested by asking for an algebraic expression and checking for algebraic equivalence, mistakes in the derivation can be caught and provided feedback for by using additional nodes in the potential response tree (PRT). In some cases, for both numerical applications and simple derivations, it was necessary to split the question in two in order to make the assignment of partial marks and personalised feedback easier and more detailed, it also allowed for easier propagation of student mistake. In such cases this would result in a in a loss of complexity of the questions in exchange for better feedback for the student. For more involved derivations, the equivalence reasoning input type allows for students to type in the entire derivation, thus allowing to check for specific steps in a student's answers, as would be done with a handwritten assignment.

In some cases, as mentioned above, it was even possible to adapt into STACK a question that asks for a sketch. In particular, Meclib can be used to ask the students to give a free-body diagram, something which is asked frequently in level 1 and 2 dynamics problems. Meclib allows for the students to easily name and place vectors. Students' free-body diagrams can then be assessed by checking whether all the vectors are present and correctly named, whether they are pointing in the correct directions, and whether their relative sizes are consistent with the situation of described in the question.

Using all these functionalities in combination we were able to adapt preexisting standard physics problems into Moodle questions, either fully in STACK, or STACK in combination with the essay type question.

Providing feedback

In addition to its suitability to pose STEM questions, STACK was also chosen because of its ability to provide personalised feedback. The goal was to create assessment which would mimic as close as possible the specificity of the feedback a student might get when submitting handwritten assessment. To do so, we chose to program most of the question's feedback after the student had submitted, so that we could match the feedback to the range of answers that the students gave.

In a physics problem there can often be many ways to get the answer wrong, therefore we do not want for the feedback to rely on our ability to guess the possible mistakes a student may make. We approach this by asking the students to upload a picture of their workings alongside their submission. The workings upload is optional. Once the students had submitted, the feedback was programmed in by implementing the following steps. If the statistics show that a question is well done by the students, nothing is added to the question beyond the existing feedback. If the statistics show that the question has mixed results, the students' answers are reviewed alongside their working, and reasonable mistakes are programmed in with partial marks and accompanying feedback. Frequently reoccurring mistakes that do not warrant partial marks are still programmed in with feedback so that the students can identify and understand their mistake. The submission is then released to the students.

There are several drawbacks to providing feedback this way: this method excludes the possibility of giving the students assessment with immediate or adaptive feedback, which is one of the strengths of computer-based assessment, and it can sometimes take a considerable amount of time to properly go through the students answers and write the associated feedback. This mitigates some of the advantages of using computer-based assessment over traditional handwritten assessment. Furthermore, students who have made unique or unusual mistakes may not get personalised feedback.

This feedback method also has many advantages, some of which minimise the effects of the disadvantages. Once a question has gone through this feedback process several times, for example if the homework is given to a class several years in a row, it can eventually be used with immediate or adaptive feedback, as it will have built up a detailed and complete feedback tree, based on real student answers. Additionally, the time investment of writing feedback with this method is reduced with every use of the question, as more possible answers are caught. Providing feedback in this manner does not require that every student has submitted their workings or to review every student's workings, as a given mistake will usually be repeated by several students. This is particularly advantageous in the case where the assignment is given to a very large cohort, as this means that reviewing only a fraction of the student answers will lead to sufficiently detailed feedback.

3. Use of STACK questions for physics assessment at the University of Glasgow

Within the school of Physics and Astronomy, STACK is used across all levels to assess the students both formatively and summatively. In this section, the manner in which STACK questions are implemented for each level will be described, with an emphasis on the implementation in level 2 physics, where it is the most widely used.

Level 1 physics

In level 1 physics, the students are given Moodle quizzes as homework, both summative and formative. Some of these homework assignments make use of STACK questions, however the questions are kept in a very simple format for the students, usually in the form of simple calculation questions.

Level 2 physics

In level 2 physics, the students' overall grade is constituted of 25% class tests, 25% skills modules (mostly laboratories), and 50% degree exams. The course is divided in eight modules, where two are run concurrently per half semester. Every half semester, the students are given a class test covering the content taught in the last two modules, with one question per module, each worth fifteen marks.

To help the students practice for the class tests, online homework assignments have been created using Moodle quizzes, made up mostly from STACK questions. These assignments are optional and do not count towards the students' final grade. They are in addition to problem sets provided for discussion in regular small tutorial groups. To best prepare the students for the class tests, the homework was designed to resemble the class tests as closely as possible. This means each homework assignment is made up of two sections, one per module, each worth fifteen marks. The breakdown of the marks is given to the students, as is given in the class test. The students are assigned these homework assignments four times per semester, giving them the opportunity to practice twice per class test. When the homework is opened, the students are given a week to complete it. Unlike in the class tests which must be completed in 50 minutes, there is no time restriction on the homework.

As described in section 2.2, the students are encouraged to submit their workings with their homework to assist with grading. Once the homework is closed, the results are reviewed to identify the parts of the homework that require additional partial marks. The students' answers for those parts are reviewed alongside the workings to identify common mistakes. Additional PRT nodes are added to the STACK questions for the identified mistakes, with the associated detailed feedback for the students. If some of the questions in the quiz required image uploads or typed answers, these are graded manually. This process is usually done within a day, allowing for the students to receive the mark and feedback for their work the day after they submitted.

Physics Summer School

The University of Glasgow hosts a physics summer school for American students. The content covered is similar to that taught in first year and second year physics. In the summer school, the students' overall grade is constituted of 5% reading tests and group tests, 15% class tests, 20% laboratories, and 60% exams. The class tests and exams have the same format, and are both taken as a Moodle quiz, taken in person, and invigilated. The quizzes are constituted of twenty multiple choice questions (MCQ), and three STACK questions, of which they need to select two. The students have 90 minutes to do the test or exam. The summer school lasts 8 weeks during which the students take six class tests and two exams.

Contrarily the use in level 2 physics, in the summer school STACK is used for summative assessment. The grading process is the same as the one described for the level 2 homework. The summer school being a very intense and fast paced learning environment, STACK allows for the students to get personalised feedback very quickly after taking the test, thus giving them the opportunity to review their mistakes and correct their understanding before the course moves on to different content.

Honours level physics

Currently, STACK is only being used in a few courses at honours level, for simple expression and derivations. Work is currently being undertaken to adapt past papers questions for all honours level courses into STACK questions. As the level of difficulty of questions is increased, the adaptation of the past papers questions into STACK questions requires more work and modifications. The equivalence reasoning input type is used more frequently than for level 2 questions, as more complex derivations are expected from students, and questions requiring long answers have to be broken up into several input fields to be able to capture all the key steps in a student's answer and to facilitate propagation of a student's mistake.

The planned use for these questions is to give them to students as homework, and to use them as a self-assessment tool for students' readiness for higher level courses.

4. Analysing the usefulness of the level 2 physics homework

The purpose of the homework given to the level 2 physics students, is to help them prepare for the class tests that they must take twice every semester. Therefore, to assess the effectiveness of the homework we aim to compare the performance and participation of the students in the homework, with their performance in the class tests. While a rigorous analysis has not yet been made, the following paragraphs detail the preliminary observations which have been made, and future plans to analyse the homework results further.

Methodology

As detailed in paragraph 3.2., every two homework assignments correspond to one class test. Therefore, the performance for every two homework assignments will be compared with the performance in the corresponding class test.

As each set of two homework assignments covers the same topics, the students' performance for each set is chosen to be measured by their average score for the set. Because there are some students who will have attempted only one, or none of the homework assignments, a lack of attempt was considered to count as a score of zero, to be able to compute an average score for each student. From the average score, a letter grade was assigned to each student for every set of two homework assignments. The letter grades range from A to D for a passing grade, anything below D was counted as a fail [UoG].

When assessing participation, the optional nature of the homework and the fact that the homework was set to submit automatically when it was due, created complications in determining what should constitute a genuine attempt by a student. Indeed, in every homework some of the submissions would be from students who had opened the homework, but not answered any of the questions, or who would only start answering one part of one question before abandoning their attempt. A genuine attempt was decided to be one where a student gets five marks or more out of thirty. This threshold was chosen arbitrarily by looking through the attempts of the students scoring extremely low, and estimating the score above which most students seemed to have genuinely attempted the homework. This is a flawed method of determining whether an attempt is genuine, made necessary by the automatic submission of the homework. A better way to do this in the future would be to enable manual submission instead of automatic submission, with the disadvantage that if a student forgets to submit at the end of a serious attempt, the attempt will not be counted as completed by Moodle, potentially skewing the data.

Results

A first surface-level observation concerning the homework was that participation was good as shown in table 1, particularly considering that this homework is purely optional and that the students have no other incentive to do it other than the potential benefit to their learning.

Table 1. Level 2 student participation in the 2023-2024 homework

Homework	Participation
Homework 1	87.1%
Homework 2	76.3%
Homework 3	67.7%
Homework 4	51.6%
Homework 5	58.6%
Homework 6	52.7%
Homework 7	59.1%
Homework 8	38.7%

Table 1 shows that participation is at its highest at the start of the academic year, suggesting higher motivation and student engagement in the first month of teaching, and particularly low for the last homework of each semester (homework 4 & 8), corresponding to times where students have a lot of deadlines for assignments counting towards their final grade or need to prepare for exams in other subjects, which they are likely prioritising over completing the homework.

In order to visualise the correspondence between student performance in the homework assignments and the class tests, the students' average grade for a set of two homework assignments, and their grade in the class test, were represented in Sankey diagram, as shown in figures 1 and 2. This was done for each set of two homework assignments and corresponding class test for the first semester. Similarly, we can visualise the correspondence between participation in the homework, regardless of the grade, and the class test results, show in figures 3 and 4.

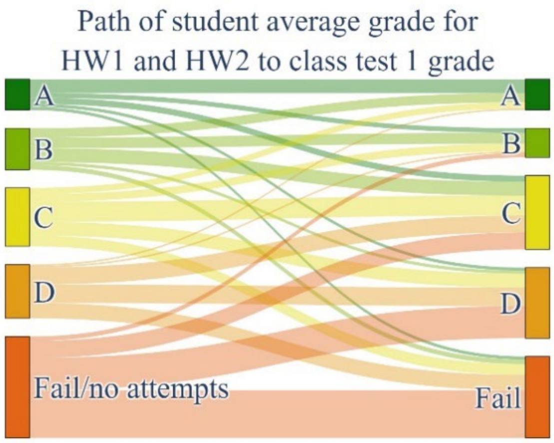


Figure 1: Path of student average grade for homework assignments 1 and 2 to class test 1 grade

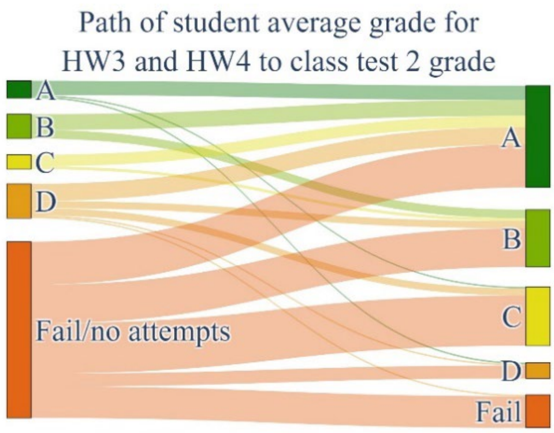


Figure 2: Path of student average grade for homework assignments 3 and 4 to class test 2 grade

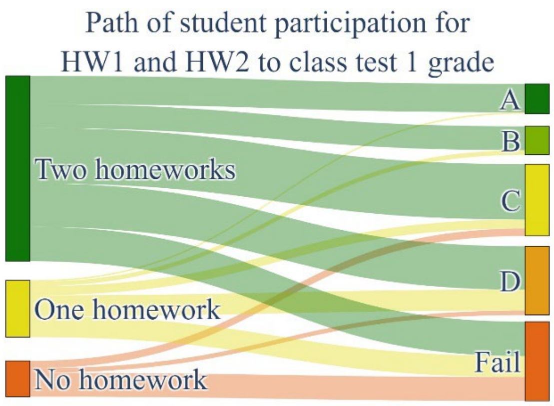


Figure 3: Path of student participation for homework assignments 1 and 2 to class test 1 grade

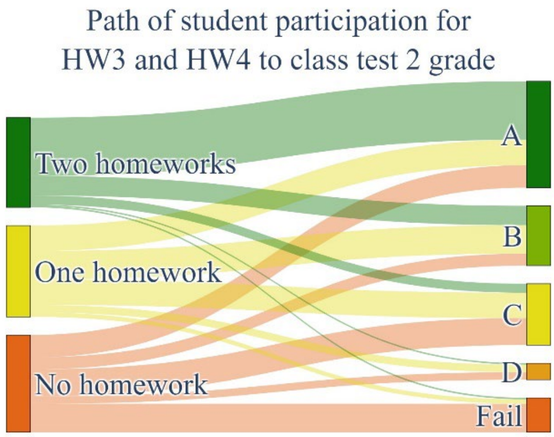


Figure 4: Path of student participation for homework assignments 3 and 4 to class test 2 grade

While these visualisations are not suitable to draw any concrete conclusions concerning the effectiveness of the homework, it allows to see if there appears to be any surprising results or major discrepancies, none of which appear to be present here. The students who attempted at least one homework assignment are unlikely to fail and the students who obtained a passing grade on the average of their homework assignments tend to pass the class test.

Future work

To analyse the effectiveness of the homework more precisely, we are planning to classify all the questions in the homework assignments and class tests using Bloom's taxonomy, and compare students' performance for each level of Bloom's taxonomy between the homework assignments and the class tests. We are hoping to determine whether the Moodle homework we created is effective in testing students on all levels of the taxonomy, or if there are levels for which the homework is better suited than others.

5. Conclusion

Over the course of our implementation of STACK to write physics problems, we have found that it is well suited to pose most type of physics questions, and that it can be complemented with other question types for the cases where it is not, thus conserving the form of the original question in the process. Additionally, we have found that by asking the students to upload their working, we can create highly personalised feedback for the students. This method is more efficient for larger student cohorts.

In particular, we have looked at the results from the homework from the level 2 physics course where STACK homework is the most widely used. Preliminary observations show that student engagement is high throughout the year, especially for assignment that is purely optional. Homework assignments based on STACK accurately differentiate the students based on their learning and the security of their knowledge.

Our work has shown that STACK, in conjunction with other online quiz types, is a suitable alternative to traditional on-paper tutorial and exam questions. It can be implemented across all level of a physics degree and can test all of the same types of skills as traditional exam questions with the added advantage giving students timely and personalised feedback.

Bibliography

[OP22] Geraldine O'Neill & Lisa Padden (2022*). Diversifying assessment methods: Barriers, benefits and enablers*. Innovations in Education and Teaching International, 59:4, 398-409, DOI: 10.1080/14703297.2021.1880462

[PMH22] Katrina Perry, Kane Meissel & Mary F. Hill (2022). *Rebooting assessment. Exploring the challenges and benefits of shifting from pen-and-paper to computer in summative assessment*. Educational Research Review, v36, https://doi.org/10.1016/j.edurev.2022.100451

[KS21] Kraska, M., & Schulz, D. (2021, June 9). *Automatic assessment of free body diagrams using STACK*. International Meeting of the STACK Community 2021. https://doi.org/10.5281/zenodo.5083692

[UoG] *Guide to the code of assessment - 2 grading student performance*, pp11-12. https://www.gla.ac.uk/media/Media_124293_smxx.pdf

Focuses: Implementation and usage of GeoGebra, JSXGraph, or other Programs; New users and authoring of questions.

Article number: 15

STACKing Further with STACK-JS

Sam Fearn*

Durham University

Abstract

This paper explores the use of STACK-JS, as a tool for enhancing STACK questions through the addition of custom JavaScript and the inclusion of existing JavaScript libraries. As an example of how STACK-JS may be used, we focus here on an example which creates a custom method of user interaction, designed to allow for a novel means of testing mathematical proof. In particular, this example was created to be used in a proof-heavy first-year undergraduate Analysis module at Durham University, emulating a type of problem often used in Computer Science courses known as Parsons Problems.

* Corresponding author: s.m.fearn@durham.ac.uk

1. Introduction

Assessment and feedback is a fundamental part of a degree in mathematics. Formative assessment may be defined as “encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (Black and Wiliam, 1998). This style of assessment is also often referred to as ‘assessment for learning’, to emphasise the idea that the assessment is in support of learning, as opposed to ‘assessment of learning’ which is designed to certify that a learner has met a set of learning objectives. Such assessments give course leaders both a snapshot into the progress of the class as a whole, as well as a chance to identify particular students who may be struggling and in need of further support. They also allow students to gauge their own level of understanding, and the feedback obtained from an assessment should enable a student to further their learning. Indeed, studies have shown that “innovations that include strengthening the practice of formative assessment produce significant and often substantial learning gains,” but moreover that “improved formative assessment helps low achievers more than other students and so reduces the range of achievement while raising achievement overall” (Black and Wiliam, 2005).

However, in order for feedback to be effective, students must be given the opportunity to make use of it (Shute, 2007). Even when staff are able to return marked work quickly, by the time students receive their feedback, they have often moved on to a new topic and the feedback they receive may seem less pertinent. Automated assessment offers the opportunity for feedback to be generated at the moment of submission, giving students the chance to immediately close the feedback cycle, putting the received feedback into practice by re-attempting a given question (Sangwin, 2013). Moreover, automated assessments allow for questions to be generated using randomised inputs, so that when a student does re-attempt a question, they cannot simply input a correct solution they have been given in previous feedback. Not only does automated assessment therefore represent an opportunity for enhancing learning, it also clearly offers workload benefits to staff and their departments, saving the need for manually marking students' submissions.

Automated assessments have traditionally been best suited to types of questions that ask a student to do some routine calculation, as is often required in first courses in calculus or linear algebra (Sangwin, 2013). However, higher level mathematics also requires students to develop their proof comprehension, and this has traditionally been an area where it is difficult to take advantage of automated assessment. Although ideas for the automated assessment of mathematical proof have been discussed in the literature previously (Bickerton & Sangwin, 2022), this remains a crucial area for the development of automated assessment tools.

While exploring these ideas in the context of a proof-heavy first-year undergraduate Analysis module, I experimented with testing students’ proof-comprehension skills using a type of problem known in the Computer Science literature as Parsons Problems (Ericson et al., 2022). Specifically, rather than asking a student to write their own proof of a given result, they were instead asked to arrange pre-defined statement blocks in order to construct a valid proof. The Department of Mathematical Sciences at Durham University uses STACK (Sangwin, 2013) for automated assessment, and at that time STACK did not natively provide a means of constructing such a question. A crude approximation was possible within our environment using an alternative tool (another Moodle question type), but was limited by requiring a single fixed answer without distractors. Since many valid mathematical proofs allow for the interchange of particular logical blocks, such as when proving the equality of sets by showing each is a subset of the other, this was not sufficient for the required use case. However, STACK does allow a means of augmenting the inbuilt capabilities through the addition of custom JavaScript, enabling advanced Parsons problems to be created. In this paper, I will discuss how STACK-JS can be used to extend STACK (Harjula, 2023), enabling advanced and custom forms of user interaction among many other possibilities. The previously mentioned Parsons problems will serve as a contextualising example throughout.

2. Using STACK-JS

STACK-JS is a tool for including custom JavaScript (and existing JavaScript libraries) within a STACK question, whilst providing security by separating the executing code from the host Virtual Learning Environment (VLE). In principle this VLE could be Moodle or ILIAS, though the rest of this paper assumes STACK is being used via Moodle. This separation is achieved through the creation of an iframe in which the JavaScript runs, with limited interaction possible between the JavaScript and the VLE session. Crucially however, STACK input fields can be updated by the JavaScript, meaning that user interactions within the iframe can result in inputs that are then evaluated by STACK.

At a high level, which we then discuss in more detail through the example of creating a Parsons problem below, the key steps for integrating new JavaScript with STACK using STACK-JS are:

1. Use `[[iframe]] and [[script]]` blocks to create the iframe, and import any required JavaScript libraries;
2. Create a (hidden) STACK input for any state data you want to communicate between STACK and the JavaScript;
3. Use the STACK helper function `stack_js.request_access_to_input(ans1,true)` to get access to the STACK input `ans1` in the JavaScript;
4. Optionally, update the STACK input after changes in the iframe.

These steps may look familiar to readers who are already familiar with writing STACK questions that utilise the graphing tool JSXGraph (Sangwin, 2018). JSXGraph is itself a JavaScript graphing library which is officially supported for use in STACK. Using JSXGraph in STACK is therefore simplified when compared to using arbitrary JavaScript, through the existence of a number of helper functions. However, the underlying structure of a STACK question using JavaScript is very similar in either case.

3. Worked example: creating a simple Parsons problem

In order to create a simple Parsons problem in STACK, we need to create a list of statements that the students should sort, as well as provide a means of interacting with the list of statements. We first consider the case where there are no distractors in the list of statements, and therefore the student simply needs to sort a given list into a correct order. A simple example question may then be as shown below in Figure 1, where the statements to be sorted are simply the numbers one to five.

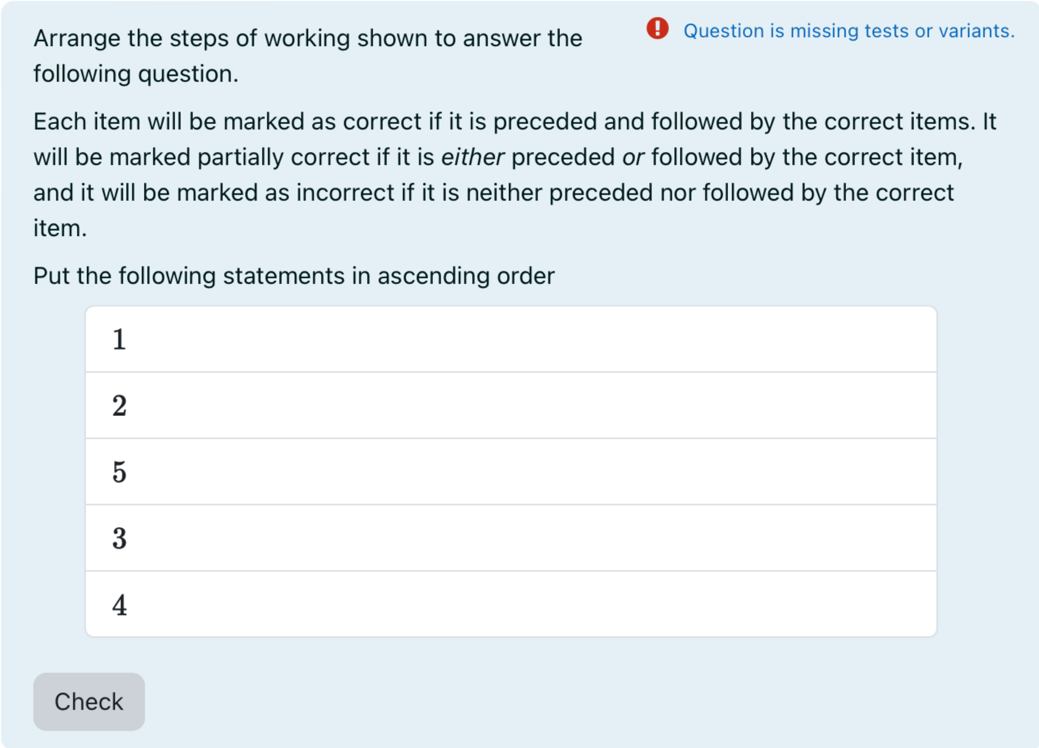


Figure 1: A simple Parsons problem using STACK-JS and SortableJS

Initialising the iframe and loading libraries

There are many JavaScript libraries one could use to create a sortable list of items, but for the purposes of this example we will use the Sortable library (Mills, 2024). We therefore create an iframe within the question text of our question, and load the STACK-JS and Sortable libraries:

```
<p>Put the following statements in ascending order</p>
[[iframe]]
[[script type="module"]]
import {stack_js} from '[[cors src="stackjsiframe.js"/]]';
import '[[cors src="Sortable.js"/]]';
```

Here, the `[[iframe]]` block is used to create an iframe within the generated HTML of the question, and the `[[script]]` block is used to define a script element within the header of this iframe. The script element is given the attribute `type="module"`, so that other JavaScript libraries can be imported in, but this element can also be the one in which we include any additional custom JavaScript we need. The `[[cors]]` blocks provide a simple way to reference the paths to our JavaScript libraries, which are stored locally on our Moodle server. The file location for such libraries is then specified relative to the `moodle/question/type/stack/corsscripts` directory. Note that later in the question text we will need to close the script and iframe blocks.

Input, State and STACK-map

All STACK questions should have at least one input, though since students are to interact with our Parsons problem by dragging to reorder the list of statements, we may not want this input to be visible. Even when not visible, this input will be used to store the state of a student's response to our problem. We can create such a hidden input in our question text as follows:

```
style="display:none">[[input:statestringinput]][[validation:statestr
inginput]]
```

The "`display:none`" style command is used to visually hide the input box from the student. There are of course serious accessibility issues with creating a question that only allows interaction through using a mouse (or touch input) to re-order list items, and so this approach may not be the most suitable. Since we use this Parsons problem merely as an example of how to use STACK-JS, we leave further discussion of this point for other work.

As implied by the given name of the input, we will store the student's answer as a string, so the input should be given type string. Since the input field itself is hidden, we should use the input options to not require or display validation, and the extra option `hideanswer` may be specified to ensure that information about the "teacher's answer" is not displayed at any point to students. Note that we have not yet given the student a way to submit their answer through this input; this will be discussed in the next section.

The statements that our students will sort are the key question variables, and so within the question variables field we initialise a list of statements. Since these statements may be long strings, potentially involving mathematics displayed using MathJax, it will be much simpler to refer to these strings wherever possible using unique short keys. A data structure consisting of keys and associated values (our statements to be sorted) can be represented in our JavaScript as a JavaScript Object. Although such a structure is not native to Maxima, STACK offers helper functions for dealing with data in this form as a so-called STACK-map. This is a nested list structure in Maxima, with the first element of the parent list being the string "`stack_map`", and key:value pairs as sublists for our keys and statements. In our example, we might therefore initialise our question variable as follows:

```
correctlist:["stack_map",["f","\\(1\\)"],["y","\\(2\\)"],["j","\\(3\
\)"],["q","\\(4\\)"],["v","\\(5\\)"]]
```

Here we have simply chosen random keys, though of course in practice one could easily write a function to create this structure in Maxima simply from a list of the strings we want to sort. We also note that as well as both our keys and statements being stored as strings in Maxima, the statements make use of inline LaTeX maths. The double backslash is required here, as the backslash character is a so-called special character, and hence requires 'escaping' with the additional backslash to ensure it is parsed properly once passed to the JavaScript. In this simple example, this STACK-map stores the statements in the unique correct order; we discuss alternative correct orders in the context of a more complicated problem later in the section 4 of this paper.

Our statements are now stored in key:value pairs in Maxima as a STACK-map, and will be handled as an object in the JavaScript. We can easily convert the data between these two forms by utilising JSON, more specifically JSON strings. It is the JSON string representation which we will store in the `statestringinput` STACK input. STACK includes a helper function for creating a JSON string from a STACK-map, which we can use as `stackjson_stringify(correctlist)`. We can similarly convert back from a JSON string `stateString` to STACK-maps using `stackjson_parse(stateString)`.

Connect the Maxima and JavaScript

We now have a Maxima STACK-map variable representing the statements we want our student to sort, a hidden input which should be a string representing the student's answer, and an iframe with a script element in which we have loaded a JavaScript library and can also write additional JavaScript. We now need to connect these separate pieces together, displaying the question statements, providing a means for students to sort the list items in the question, and have this update the hidden string input so we can then check the student's answer.

We first tell our JavaScript about the STACK input field we have created:

```
var stateStorepromiseinput =
stack_js.request_access_to_input("statestringinput", false);
```

This uses a JavaScript Promise to asynchronously return the id of a hidden HTML input which is created inside the iframe, and whose contents are synchronised (on change events, unless the additional boolean option to `stack_js.request_access_to_input` is set to `true`, in which case also on input events) with our hidden STACK input. This means that if our JavaScript sets the contents of the hidden HTML input in the iframe to a value representing the student's answer, our STACK input will be updated to contain the same string, which can then be submitted and checked (as discussed in the next section).

Once our iframe has created its hidden input, the JavaScript promise resolves to the id of this element. We can then load some state into this input and display the current state.

```
stateStorepromiseinput.then((stateid) => {
    let stateStore = document.getElementById(stateid);
    var state;
    // Load existing state, or initialise from a default
```

Here, `stateStore` is the hidden input element in the iframe. When a student first attempts our problem, we simply want to load a default state by randomising the order of our statements. However, if a student returns to view the question, after checking their answer for example, we will want to use the state which has been stored in our STACK input, `statestringinput`. As discussed in the previous section, `statestringinput` will be a JSON string, and so when we load this into our JavaScript, we will want to parse this JSON string into a JavaScript object. The default state will similarly need to parse a JSON string representation of our initial list of statements `correctlist`, before randomising the order of the key:value pairs (which for brevity we omit the details of below).

```
// If we already have a stored state in the statestringinput input,
then we use this state
if ( stateStore.value && stateStore.value != ''){
    state = JSON.parse(stateStore.value);
}
// otherwise our state is loaded from the correct list given as a
Maxima variable
else {
    var stateCorrect = JSON.parse({#
stackjson_stringify(correctlist) #});
    // state = shuffle(stateCorrect)
}
```

Creating a Sortable list

The preceding steps are quite generic, and will cover many possible use cases for STACK-JS. What follows in this section is specific to our example of a Parsons problem, and is included as an example of how one might use the state data within the JavaScript.

With the problem data now available in the JavaScript, we can create the visual display of our statements as an HTML list, and use the Sortable library to both make this a manipulatable element and to update our stored state whenever the student re-orders the statements. We should firstly create an empty HTML list within our iframe, which we can then populate using our JavaScript. We create the empty list after the closing of the script block, but before the closing of the iframe block as follows:

```
<div class="container"><div class="row">
<ul class="list-group col" id="correctListHTML"></ul>
</div></div>
```

Once we've loaded our state, we can then populate this empty list from inside our JavaScript:

```
let correctListHTML = document.getElementById("correctListHTML");
for (const key in state) {
    let li = document.createElement("li");
    li.innerText = state[key];
    li.setAttribute("data-id",key);
    li.className = "list-group-item";
    correctList.appendChild(li);
};
```

Here, we set the key as the value of the `data-id` attribute, as the Sortable library provides a method for returning an array of the `data-id` attributes when the list is sorted by the student. The HTML class is specified to allow for styling using Bootstrap, though we do not discuss this point further.

Finally, we use the Sortable library to add drag-and-drop interactivity to our HTML list:

```
var sortableMainList = Sortable.create(correctListHTML, {
    onSort: (evt) => {
        updateState(sortableMainList);
    },
});
```

This adds an event handler to the list, which calls the following `updateState` function every time the list is sorted:

```
function updateState(sortedCorrect) {
    stateStorepromiseinput.then((stateid) => {
        const newState = {};
        sortedCorrect.toArray().forEach((mykey) => {
            if (state[mykey]) {newState[mykey] = state[mykey]};
        });

        let stateStore = document.getElementById(stateid);
        stateStore.value = JSON.stringify(newState);
        stateStore.dispatchEvent(new Event('change'));
        state = newState;
```

```
    });
}
```

This creates an array of the `data-id` attributes of the list items, representing the order the student has sorted the original keys into, and then creates a `newState` object consisting of the keys and corresponding statements in this new order. This is then converted to a JSON string using `JSON.stringify(newState)`, which is in turn stored in the hidden input in the iframe. Since this iframe input has been configured to be synchronised with our hidden STACK input, we therefore have a JSON string representing the sorted list in our STACK input.

Marking the attempt

In this simple Parsons problem, the correctness of a student's answer may be determined by comparing the student's ordered list of keys from the `statestringinput` STACK input with the ordered list of keys specified in the initial question variables within the `correctlist` variable. Since the marking of the answer is done within STACK, using variables configured in the feedback variables field of a potential response tree (PRT), we should convert the JSON string from `statestringinput` back to a STACK-map for analysis with Maxima. As mentioned above in the section on inputs and STACK-maps, we can create a Maxima variable `stateString` using the helper function `stateString:stackjson_parse(statestringinput)`. Since we only need to compare the keys from the STACK-maps, we can extract just these keys into two respective lists as follows:

```
correctkeys:stackmap_keys(correctlist);
sakeys:stackmap_keys(stateString);
```

There are many ways one could produce a score, and indeed feedback, given these two lists. One possible algorithm, which we present here without the explicit Maxima code for brevity, is:

- Let n be the number of items the student is sorting, so in our simple example n = 5.
- For the first item in the student's list, give a score of $\frac{1}{2n}$ if the key for this item is also the first key in the teacher's list, otherwise set the score to 0.
- For the first item in the student's list, add a score of $\frac{1}{2n}$ if the second key is the second key in the teacher's list.
- For items $k \in \{2, \ldots, n-1\}$ in the student's list: check whether key $k-1$ in the student's list is the same as key $k-1$ in the teacher's list, if so add a score of $\frac{1}{2n}$. Similarly, add a score of $\frac{1}{2n}$ if key $k+1$ matches in both lists.
- For the final item in the student's list, add a score of $\frac{1}{2n}$ if the preceding key matches in both lists. Similarly, add a score of $\frac{1}{2n}$ if this is also the final key in the teacher's list.

This algorithm gives partial credit for items being locally correct when compared to their neighbours, even if a student has the wrong absolute order (due to an incorrect first item for example). If a student submitted the answer as shown above in Figure 1 for example, they would score 0.5/1.0, since the numbers 1 and 2 are in the absolute correct position (scoring a total of 0.3), and the numbers 3 and 4 are adjacent items in the correct order (scoring an additional 0.2) despite being in the wrong absolute position. We briefly return to the case of their being multiple correct answers in the following section.

4. A more complicated Parsons problem

While the example discussed in detail in the previous section demonstrates how to use STACK-JS for a simple Parsons problem, we can make some additions which greatly improve how useful this would be for assessing a mathematical proof. Here we present only the broad steps required for two important additions we might make.

The first limitation of our simple example is that it doesn't allow for distractors — items which should be separated out of the correct list. As an example, a distractor could be used to test a student's understanding of logical qualifiers in a proof, with statements that differ only by whether "there exists an element of the set *X*", or "for all elements of the set *X*". In order to allow for this, we can add a second list to our Parsons problems, with one list for indicating the correct statements in the correct order, and the second for indicating the statements which should not be included in the proof (irrespective of order). We start with all statements displayed in the 'incorrect' list and ask the student to filter the valid proof steps into the 'correct' list, in the correct order. Although we would still only need to mark the 'correct' list, we would want to preserve the order of the items in both lists visually when a student checks their answer. We can achieve this by extending our state variable such that the corresponding JavaScript object contains both a `correctlist` and an `incorrectlist`, both of which are themselves JavaScript objects of key:value pairs similar to our simple example.

Secondly, our simple example supports only a unique correct answer. A proof involves showing two sets are equal, by showing each set contains the other as a subset, has the obvious freedom to demonstrate the subset inclusions in either order. Such a Parsons problem would therefore not have a unique solution. We can support such cases by allowing the question author to specify a list of alternative correct orders for the keys. When marking the student's answer, we can then apply a modified version of the algorithm presented previously. In particular, for each key in the student's answer we can consider the following keys in all alternative answers, and give a mark if the following key in the students answer matches any of the valid following keys. We then do similarly for the preceding keys.

An example question demonstrating these features is shown below in Figure 2. In this image, the student has already constructed a correct answer to the question. If the student were to move the fifth and final statement from their 'correct' left-hand list into the second position in this list (with statements two to four simply moving down in fixed order), they would still score full marks for this question, as this represents an alternative correct order as specified by the question author.

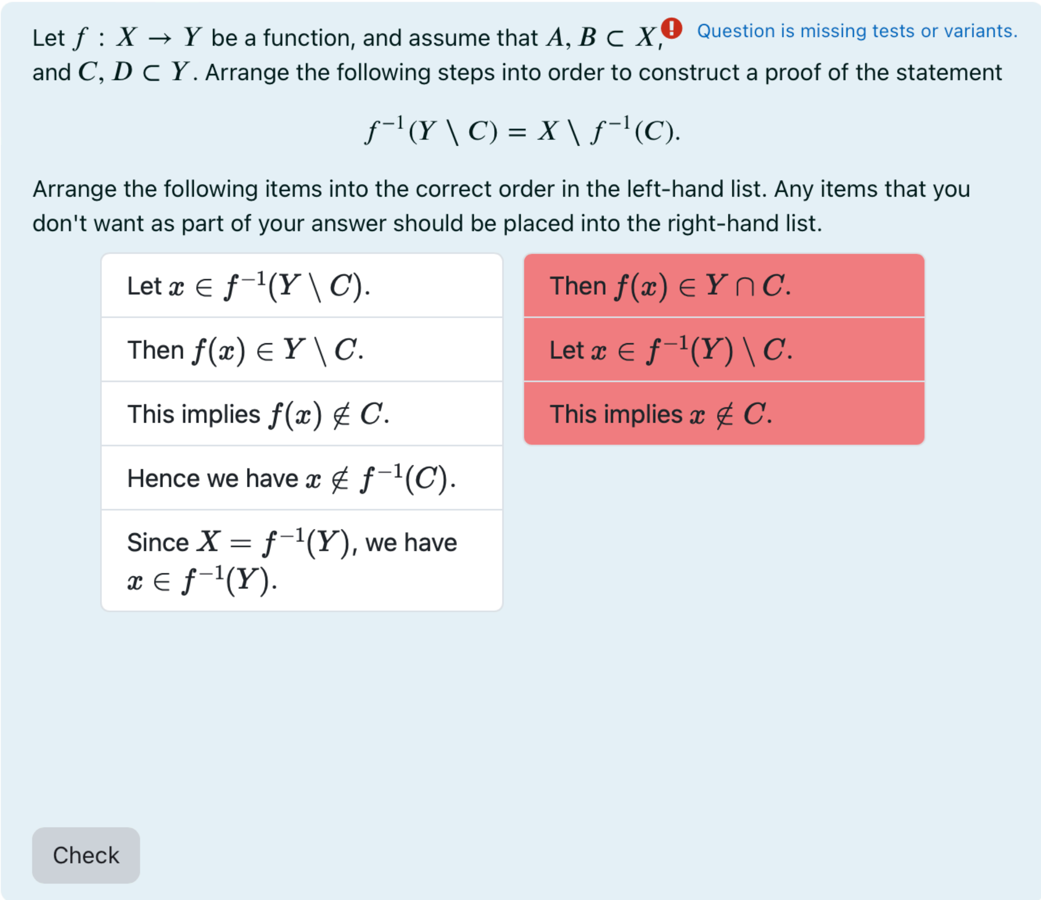


Figure 2: A more complex Parsons problem using STACK-JS and SortableJS, demonstrating the possibility for distractors in the problem.

5. Summary

In this paper we have presented the key steps for integrating custom JavaScript into STACK questions using STACK-JS. This enables possibilities such as new input methods for students to use, or advanced visualisation methods, among others. Here, we presented an example of the former use case, creating drag-and-drop Parsons problems (Ericson et al., 2022) suitable for testing students' proof-comprehension skills. Examples of the latter use case might include using the D3 library (Bostock, 2024) for data visualisation, or VisualPDE (Walker et al., 2023) for the visualisation and exploration of 1d and 2d PDEs and their solutions.

Following conversations with the core developers, STACK now natively supports drag-and-drop Parsons problems (Sangwin, 2023), building on the ideas presented in this paper. We would therefore not recommend that question authors use the specific implementation of Parsons problems presented here, in favour of using the native implementation. However, we hope this example nevertheless serves as an effective demonstration of how to use STACK-JS to add advanced features to STACK questions.

Bibliography

Bickerton, R. and Sangwin, C. (2022). *Practical on- line assessment of mathematical proof.* International Journal of Mathematical Education in Science and Technology, 53(10), pp. 2637–2660. https://doi.org/10.1080/0020739X.2021.1896813

Black, P. and Wiliam, D. (1998). *Assessment and classroom learning*. Assessment in Education: principles, policy & practice, 5(1), pp. 7–74. https://doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (2010). *Inside the Black Box: Raising Standards through Classroom Assessment*. Phi Delta Kappan, 92(1), pp. 81-90. https://doi.org/10.1177/003172171009200119

Bostock, M. (2024). D3. GitHub. https://github.com/d3/d3

Ericson, B. J., Denny, P., Prather, J., Duran, R., Hellas, A., Leinonen, J., Miller, C. S., Morrison, B. B., Pearce, J. L., and Rodger, S. H. (2022). *Parsons problems and beyond: Systematic literature review and empirical study designs*. Proceedings of the 2022 Working Group Reports on Innovation and Technology in Computer Science Education, pp. 191–234. https://doi.org/10.1145/3571785.3574127

Harjula, M. (2023). *STACK-JS*. STACK Docs. https://docs.stack-assessment.org/en/Developer/STACK-JS/

Mills, O. (2024). Sortable. GitHub. https://github.com/SortableJS/Sortable

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

Sangwin, C. (2018). JSXGraph. STACK Docs. https://docs.stack-assessment.org/en/Authoring/JSXGraph/

Sangwin, C. (2023). *Support for proof in STACK*. STACK Docs. https://docs.stack-assessment.org/en/Proof/

Shute, V. (2007). The future of assessment: Shaping teaching and learning. *Tensions, trends, tools, and technologies: Time for an educational sea change*, pp. 139–187.

Walker, B.J., Townsend, A.K., Chudasama, A.K., and Krause, A. L. (2023). *VisualPDE: Rapid Interactive Simulations of Partial Differential Equations*. Bulletin of Mathematical Biology 85(113). https://doi.org/10.1007/s11538-023-01218-4

Information on the author

Sam Fearn is an Associate Professor (Education) in the department of Mathematical Sciences at Durham University. Before starting this role in 2020, he worked as a Teaching Fellow in the same department, following the completion of his Ph.D. in 2018.

Focuses: New users and authoring of questions; STACK in teaching or exams.

Article number: 16

Using STACK Beyond Mathematics

Maciej Tadeusz Matuszewski*

Durham University

Abstract

STACK is often seen primarily as a tool to aid the teaching and assessment of Mathematics. However, it has the potential to be used in a far wider range of quantitative subjects. This paper will explore the practicalities of extending the use of STACK to a wider audience, including the training of new staff to use STACK. The experience of the Durham University Mathematical Sciences Department will be used as the primary example – discussing the use of STACK in modules with Physics content in Mathematical Sciences, exploring the engagement of the Department's STACK team with other Departments within the Natural Sciences Faculty, and reviewing efforts to introduce optional STACK workshops in general lecture training courses. The paper will outline current ideas for best practice, and open discussion for new approaches.

* Corresponding author: m.t.matuszewski@durham.ac.uk

1. Overview of Implementation of STACK In Durham Maths

The University of Durham is a research based institution in the North East of England. The Mathematical Sciences is a large department, each year welcoming around 250 new undergraduate students studying Mathematics degrees (on both three and four year courses), and close to 100 postgraduates. In addition to these students, the department also provides service modules (introductory courses in mathematics) to students from other departments (such as Physics) in the Faculty of Natural Sciences, and offers courses to students on the joint Natural Sciences degree. This results in a large teaching workload, with the largest first year courses having close to 500 students. With most first year modules having homework assessments every week and most second year modules every two weeks, the marking workload was particularly heavy.

Virtual Learning Environments such as Blackboard do have a number of native automated e-assessment solutions, such as multiple choice questions. These are easy to use for both students and teachers, and have significant pedagogical uses (Huntley et al., 2009). However, these methods can struggle in assessing students' methods for solving mathematical problems, and providing adequate feedback, which is an important part of assessment (Shute, 2007). If questions are not carefully designed, it can also become possible for students to solve them using methods not anticipated by teachers, which do not help with the desired learning outcomes (Sangwin & Jones, 2016). Therefore, educational gains can often be less than with other forms of assessment (Attali, 2015).

In 2019, the Department therefore introduced a trial of using automated e-assessment using the similar STACK and Numbas systems. With the ability to assess freely typed numerical and algebraic answers, these systems replicate a good deal of the flexibility of manually marked assessment, and allow students to receive detailed automated feedback instantly (Sangwin, 2013). This pilot was judged a success, with good student and staff satisfaction. A survey of students in the Single Mathematics A module (a service module for other departments in the Faculty) for the 2021-22 academic year showed that only 13% of them preferred manually marked assessment to automated e-assessment (with the remaining students either having no preference or preferring both equally). Student attainment in final exams for modules within the trial remained roughly consistent before and after e-assessment was introduced.

Automated e-assessment, now focussing primarily on the STACK system due to its greater flexibility, was therefore introduced more widely throughout the Department, including to modules at higher levels and including more pure mathematics (Matuszewski, 2023). In the 2023-2024 academic year, the use of e-assessment resulted in a total saving of close to 10,000 student homework scripts no longer having to be marked manually.

The University runs the Blackboard Ultra Virtual Learning Environment, which does not have a STACK plug-in. Therefore, the Mathematical Sciences Department runs its own Moodle server, which does allow for quizzes with the STACK question type, which is linked to Blackboard Ultra via an LTI connection. This allows students to access their STACK quizzes on Moodle seamlessly via a single link on Ultra, with no further log in required. This solution allows for greater flexibility, with Departmental staff having full control of the Moodle and STACK installations but does require Departmental staff to be fully responsible for the server. This is a significant, time commitment, though manageable in the case of only hosting quizzes for Mathematical Sciences modules.

With its successes within Mathematical Sciences, there is therefore a good argument for automated e-assessment with STACK to be introduced further throughout the University. As the Mathematical Sciences Department already uses STACK in first year service courses for other Departments, and in more physics based modules for maths students (such as Special Relativity and Electromagnetism), it seems clear that automated e-assessment is suitable for more applied material, and is appreciated by students outside of mathematics. However, there are some practical considerations, which will be explored in further sections.

2. Staff Training

STACK is a system that requires a significant amount of training before it can be used to its full advantage. In the pilot stage of STACK's usage within the department, this training was provided in an ad-hoc manner.

In initial years questions were authored by a small group of members of staff. They received initial training via attending external workshops, such as those provided by Edinburgh University or by the EAMS conference. These workshops were very good at producing introductory material but lacked flexibility, both in terms of when they were scheduled, and the ability to cover specific material that would be useful for the Department. However, the staff chosen to do the initial question writing all had good general technical computational skills, so were able to develop their knowledge further via both independent work (reviewing the STACK documentation and other questions), and by discussion with other more experienced colleagues. These initial members of staff became an informal core STACK question authoring team, who then further assisted any new staff.

As the use of STACK was expanded to new modules, not all module leaders were equally comfortable with the technical aspects of STACK. A dual approach was therefore taken. Module leaders who did not have time to write their own STACK questions, or did not feel comfortable doing so, were paired with a member of the core question authoring team – with the module leader providing written questions, which were then turned into STACK questions by their partner; with the two of them then reviewing the performance of the questions together. For more advanced modules, where writing questions was more difficult, larger teams worked together to write questions (Matuszewski, 2023).

However, many new module leaders did wish to write their own STACK questions, but required more support. The core question setting team created their own training workshops. These were initially informal, and organised solely within the department. More recently, a workshop has been scheduled with the Durham Centre for Academic Development (DCAD – the department responsible for supporting developing teaching and learning skills of academic staff in all other academic departments). Workshops scheduled through DCAD are available to all staff within the university. They can usually be used by staff towards their learning requirements for the Postgraduate Certificate in Academic Practice (PGCAP) qualification. In order to fit the requirements of this, the Mathematical Sciences staff planning the workshop had to make sure that the workshop was pitched as broadly as possible, and that it particularly focused on PGCAP assessment criteria (such as the effective use of feedback). It is hoped that by making this workshop broader, it will attract a wider audience, who might otherwise not have considered automated e-assessment, and introduce them to the topic.

3. Existing Use of STACK in Modules with Physics Content

A number of applied modules with Mathematical Sciences already have a heavy focus on other sciences and engineering areas. Two useful modules to review would be the first year Dynamics module, and the second year Special Relativity and Electromagnetism module.

First year Dynamics primarily focuses on teaching students how to use differential equations and other mathematical methods to solve physical dynamical systems. The mathematics involved is relatively simple for STACK to handle, however, clever design is required to ensure that questions appear in a way that is natural for a physical problem.

Figure 1 shows an example of a question from this module, which firstly asks students to find the moment of inertia of a certain system, and then use that to find the total energy of the system. The design of the question allows students to enter both the kinetic and potential energy in one field, and, as shown in figure 2, the question is able to check individually if each of those terms is correct. This is achieved by using the Maxima ‘args’ function in the question’s Feedback variables field. This Maxima function can take any summation of terms as an argument, and returns each individual term within that summation as a separate element in a list. Each individual term (and also each pair of terms, as the potential energy can be equally well expressed as one or two terms) is then compared to whether it is algebraically equivalent to each of the correct potential and kinetic energy terms.

Furthermore, a careful construction of potential response trees (Matuszewski, 2021) allows for incorrect answers to the first part of the question be used by the STACK system to calculate what the student should expect to get if they then used the appropriate method to calculate the second part of the question. The student can therefore get appropriate partial credit and feedback. This is demonstrated in figures 1 and 2. This demonstrates the flexibility of use of STACK in complicated applied problems.

A uniform thin straight rod has length L and mass M. It is free to rotate in a vertical plane, about a horizontal axis a distance $\frac{2L}{5}$ from one end of the rod. What is its moment of inertia I?

$I =$ (7*L^2*M)/75

Your last answer was interpreted as follows:

$$\frac{7\,L^2\,M}{75}$$

The variables found in your answer were: $[L, M]$

Let θ denote the angle by which it deviates from its stable equilibrium position.

Write down an expression for the total energy E of the rod, in terms of M, L, θ, $\dot{\theta}$ and the gravitational acceleration g. You may type θ as `theta`. Please write $\dot{\theta}$ (that is, the derivative of θ with respect to time) as θ_t, which may be typed as `theta_t`. Ensure that $E = 0$ when the rod is stationary in equilibrium.

$E =$ (7*L^2*M*theta_t^2)/150+(L*M*g*(1-cos(theta)))/10

Your last answer was interpreted as follows:

$$\frac{7\,L^2\,M\,\theta_t^2}{150} + \frac{L\,M\,g\,(1-\cos(\theta))}{10}$$

The variables found in your answer were: $[L, M, g, \theta, \theta_t]$

Fig. 1: A first year dynamics STACK question with a partially incorrect answer submitted

Your answer is partially correct.

Hints:

Your answer for I is incorrect, but it is written in terms of the right variables.

Marks for this submission: 0.00/0.25.

Your answer for E contains a correct expression for potential energy.

In your answer for E you appear to have carried over an error from the previous part of the question when calculating the kinetic energy.

Marks for this submission: 0.38/0.50.

Fig. 2: Specific feedback to the above year one dynamics STACK question with a partially incorrect answer submitted

The STACK quizzes for the second year Special Relativity and Electromagnetism module can test similarly advanced applied mathematical concepts. The integral and vector manipulation problems presented to students in this module can easily be handled by STACK. The questions, however, also ask students to consider numerical answers with units. This can be done simply by having the units placed as fixed text outside the answer box, however STACK also has answer tests which allow students to enter units within the answer box, as shown in figure 3. These tests allow for the question setter to specify a required level of precision of the answer. STACK therefore allows for a sophisticated treatment of numerical answers in a physical system.

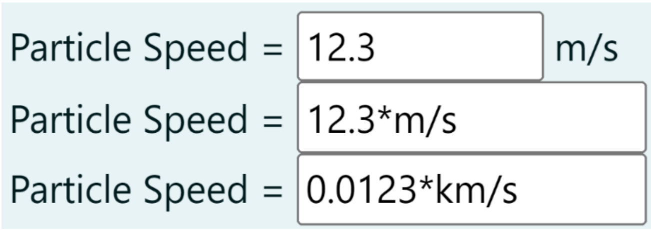


Fig. 3: Three different ways in which a correct solution to a question can be entered. The first example is of an input being tested using the 'AlgEquiv' answer test, the second and third examples are of an input being tested using the 'UnitsAbsolute' answer test. The same test can accept equivalent answers with different units as both being correct.

4. Prospects for Further Expansion of STACK Beyond Mathematics in Durham

STACK questions have been shown to be successful in assessing the type of applied mathematics common in other Natural Sciences departments in Durham. STACK questions have also been shown to be well received by students within other departments studying modules run by Mathematical Sciences. Therefore, it seems that a move to introduce assessment using quizzes with STACK questions to modules run by other departments within the University would be appropriate, following the example of other institutions which use STACK assessment beyond mathematics departments (Stetzka & Thevanesan, 2021).

However, there remain some practical considerations, though none of them insurmountable. First of all, is the novelty of STACK to many academics in other departments. When they have a workable system for assessment, it can be difficult to introduce a new system, even when it has clear advantages. Within Mathematical Sciences, STACK was introduced thanks to the advocacy of a small number of pedagogically focused staff members, and the same may be done in other departments. This has initially been attempted by informal conversations between current STACK users within Mathematical Sciences and their contacts at other departments. Additionally, members of the Mathematical Sciences department have given talks on STACK at cross-departmental workshops and conferences within the University. This has promoted interest in STACK, and it is hoped that it will drive attendance at the aforementioned upcoming DCAD STACK workshop, which will be able to fully showcase the advantages of STACK.

Holding the cross-departmental workshop will also address the initial issues of training new members of staff to author questions with STACK. It is hoped that if this workshop is successful, more can be scheduled, both within the DCAD workshop structure and by individual departments. As more members of staff across the University are trained, this network of STACK users will be able to provide additional support to new users, and the wider implementation of STACK will be able to be used to persuade the University to provide more training resources centrally.

Finally, the current arrangement of hosting a Moodle server for STACK assessment on a machine physically within the Mathematical Sciences Department and managed by Mathematical Sciences staff is not scalable. Significantly larger numbers of users would place a greater load on the server, which would require more maintenance, which the limited number of technical staff within Mathematical Sciences would not be able to provide at the required standard of reliability. Therefore, the University's Computer and Information Services (CIS) have agreed to create a virtualised Moodle server, that will be able to be used by the entire University. Their greater technical resources will be able to provide a more reliable service for a larger number of users. While this will mean less control over the server (for example over upgrades to the server), similar approaches have been used successfully by other institutions using STACK. CIS was persuaded to allow the implementation of another Virtual Learning Environment in addition to the centrally approved Blackboard Ultra as this was argued to be used to provide a very targeted service (quizzes using STACK questions) which had proved very successful, rather than replacing Blackboard Ultra. A good professional relationship between members of staff within Mathematical Sciences and CIS has made this process smoother.

5. Conclusion

The implementation of STACK within the Mathematical Sciences Department at Durham University has shown great success. From a relatively small pilot programme, STACK quizzes are now widely used for lower level assessment within the Department, with work ongoing to expand it's use further. Careful implementation of STACK has, over the long term, saved well over a thousand hours of staff time per year in reduced marking time, while allowing students to receive a comparable level of feedback to before, with far less delay. STACK has been introduced to a number of different types of maths module, from pure mathematics to mathematical physics.

This widespread success suggests that STACK would be of benefit to other departments, in particular within the Natural Sciences Faculty. While this is not without difficulty, it appears that there is significant appetite for a wider implementation of STACK, and ways to achieve this have been well thought through, and their implementation has been begun. Similar approaches should be suitable at other institutions.

Bibliography

Attali, Y. (2015). Effects of multiple-try feedback and question type during mathematics problem solving on performance in similar problems. *Computers and Education/Computers & Education*, 86, 260–267. https://doi.org/10.1016/j.compedu.2015.08.011

Huntley, B., Engelbrecht, J., & Harding, A. (2009). Can multiple choice questions be successfully used as an assessment format in undergraduate mathematics? *Pythagoras*, 0(69). https://doi.org/10.4102/pythagoras.v0i69.41

Matuszewski, Maciej. (2021). Partial Credit For Complicated Algebraic Expressions. *Contributions to the 4th International Meeting of the STACK Community*. https://doi.org/10.5281/zenodo.5036073

Matuszewski, Maciej. (2023). Applying STACK to Pure Maths an Advanced Courses. *Contributions to the 6th International Meeting of the STACK Community*. https://doi.org/10.5281/zenodo.8032289

Sangwin, C. J. (2013). *Computer Aided Assessment of Mathematics*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199660353.001.0001

Sangwin, C. J., & Jones, I. (2016). Asymmetry in student achievement on multiple-choice and constructed-response items in reversible mathematics processes. *Educational Studies in Mathematics*, 94(2), 205–222. https://doi.org/10.1007/s10649-016-9725-4

Shute, V. J. (2007). FOCUS ON FORMATIVE FEEDBACK. *ETS Research Report Series, 2007*(1), i–47. https://doi.org/10.1002/j.2333-8504.2007.tb02053.x

Stetzka, R., & Thevanesan, E. (2021). 3 Years of STACK An experience report on the implementation and use of formative randomized online tests. *Contributions to the 4th International Meeting of the STACK Community.* https://doi.org/10.5281/zenodo.5036080

Information on the author

Maciej Tadeusz Matuszewski

A Technical Manager in Durham University's Mathematical Science Department, Maciej Matuszewski is also member of the STACK Professionals' Network and the STACK Advisory Board, and will be chair of the 2025 international STACK conference.