Katja Derr, Miriam Weigel (2020). University preparation with STACK: evaluation of different question types in a self-study environment. In: Contributions to the 3rd International STACK Conference 2020. TTK University of Applied Sciences: Tallinn, Estonia. DOI: 10.5281/zenodo.3946165

University preparation with STACK: evaluation of different question types in a self-study environment

Katja Derr[1], Miriam Weigel[1]

Abstract: For the joint project 'optes', a web-based pre-course in mathematics was developed, addressing knowledge gaps in basic mathematics. The course concept comprised a diagnostic pre-test, learning modules, animations, exercises, and a final self-assessment. STACK questions were used in training sessions, each question providing feedback and detailed sample solutions. While the flexibility of this question type was found highly useful for training purposes, there were some concerns if students in a self-study environment would have difficulties with the input syntax AsciiMath. An evaluation study carried out by one of the optes partner universities focused on this issue. The results suggest that STACK questions are highly suitable for self-study learning environments and can be combined with other question types. Based on these experiences some suggestions for the design of web-based courses are made.

Keywords: mathematics, study preparation, e-learning, transition phase.

1 Introduction

One of the major challenges in tertiary mathematics education is the high heterogeneity of entering students' knowledge levels. Lacking confidence and routine in basic computations, many students struggle to meet the demands of their first year mathematics courses. Universities try to address this issue by providing preparatory courses in mathematics, but the growing demand for additional support challenges faculties and staff. One approach to this issue is to share mathematics learning material between universities. For the joint project optes[2], seven German universities and higher education institutions collaborated to design a web-based pre-course. Funded by the BMBF (German Federal Ministry of Education and Research), the project goals demand that all course contents and technical developments are Open Educational Resources, thus can be shared and adapted not only by project partners but by any third party that acknowledges the CreativeCommons BY-SA license[3]. The technical basis was ILIAS, an Open Source Learning Management System (LMS). During the development phase, each project

[1] DHBW Mannheim, Cooperative State University Baden-Wuerttemberg Mannheim, ZeMath, Coblitzallee 1-9, 68163 Mannheim, katja.derr@dhbw-mannheim.de, Miriam.weigel@dhbw-mannheim.de, https://www.mannheim.dhbw.de

[2] optes stands for "Optimierung der Selbststudiumsphase" (optimisation of the self-study phase)

[3] CC BY SA = Attribution-ShareAlike (since 2016: 4.0; before 2016: 3.0)

partner focused on a different aspect of the pre-course, from learning contents to web-based tools to e-tutoring concepts or modules on ‘learning how to learn’. The team at Cooperative State University Baden-Württemberg Mannheim (DHBW Mannheim) was responsible for the mathematical learning contents. Over a period of eight years, these materials were successively developed and integrated into the technical faculty’s study preparation. Based on quantitative and qualitative evaluations, each year a revised and improved version was launched - and re-evaluated. The final product provides self-tests and adaptive trainings for self-diagnosis and -monitoring, learning modules with exercises as well as 2D and 3D animations.

The quantitative evaluations carried out at DHBW Mannheim suggested a strong relation between learning activities like answering problems and learning gains in the pre-course. Particularly students with large knowledge gaps in mathematics were only able to benefit from pre-course participation if they repeatedly submitted self-tests and trainings ([DHA18]). Thus a focus was set on the development of large question pools that enable students to practise their basic skills in mathematics.

In its standard installation, the LMS ILIAS offers the usual question types, like MC (multiple-choice), Cloze (input of numbers), or Drag&Drop. For many applications and course contents these items are completely sufficient ([Ma99], [SJ17]) but for a larger variety and for more sophisticated questions a Computer Algebra System (CAA) was needed. Based on an analysis of existing tools and plug-ins, STACK was chosen as the most promising Open Source approach, particularly because it allows participants to enter mathematical expressions via AsciiMath code into the answer field (see also [WDH19]). STACK has been widely used in tertiary education, but in the optes community there were some concerns that the input of syntax might put too much cognitive load on learners in a preparatory course, designed for presumably inexperienced (e-) learners ([Sw05]). On the other hand, secondary students aiming at a technical degree are likely to be familiar with AsciiMath from using graphing calculators.

Based on these considerations, it was decided to more deeply analyse pre-course participants’ interactions with STACK questions, with a special focus on self-study scenarios, meaning that students should be able to answer STACK items without the support of tutors or staff.

2 Pre-course design

Prospective students may self-enrol for the pre-course on DHBW’s e-learning platform, starting in June. At the beginning of the course, learners take the diagnostic pre-test, a two-hour self-test comprising 60 items on ten topics. Based on their results in this test, students get a list of learning recommendations that are linked to the mathematics courses. The core curriculum covers six courses: 1 Arithmetic, 2 Equations, 3 Powers, roots, and logarithms, 4 Functions, 5 Geometry, 6 Trigonometry, plus ‘Basics’, a course on mathematical expressions, symbols, and units. The extended curriculum addresses topics like vectors and linear algebra, continuous functions, calculus, and logic for computer science students. The curriculum is based on the German cosh catalogue ([co14]), which is comparable to the SEFI curriculum suggestions ([Se13]).

Each of the mathematics courses addresses three to four learning goals, and for each learning goal students can find a learning module and a training. Just like textbooks, the modules provide the theoretical background, enriched by visualisations, animations[4], and short exercises. In addition, there is a training for each learning goal, consisting of 10-15 questions. The platform also provides a comprehensive mathematical glossary and a simple online calculator widget[5]. At the end of each course, students can take a final test that certifies their achievement. At the beginning of the semester, students are invited to take a final test which is parallelised with the diagnostic pre-test.

2.1 Question types

For the diagnostic test, as well as the final test, only Standard ILIAS items, like MC or Cloze questions were employed. These item types are technically easy to use and thus no barrier for inexperienced e-learners. Small exercises integrated into the learning modules were based on ILIAS standard question types, as well, whereas STACK items were added to trainings. Trainings aim at practising mathematical skills, thus all items in a training address only one learning goal. Students are free to retry a training session as often as they like.

In 2017 and 2018, the functionality and usability of a set of STACK items had been tested with first year computer science students and small groups of pre-course participants. For the 2019 pre-course, 36 STACK items were developed for the first four courses of the core curriculum and implemented into 13 trainings (=13 learning goals, see Tab. 1).

	Training 1	Training 2	Training 3	Training 4
Arithmetic	9 (6|3)	11 (5|6)	11 (7|4)	
Equations	9 (5|4)	9 (7|2)	8 (7|1)	
Powers, roots, logarithms	9 (6|3)	10 (8|2)	12 (9|3)	
Functions	10 (7|3)	8 (5|3)	5 (4|1)	9 (8|1)

Tab. 1: Number of items per training (standard ILIAS items | STACK items) in the 2019 pre-course

Fig. 1 shows an example of a training question. By clicking on the checkmark next to the input box, students could have their answers validated by the system. By clicking on the “Request feedback” button, students were informed if their input was correct (or partially correct), and they received a sample solution, and, if applicable, suggestions on the type

[4] Animated content was integrated via GeoGebra, see www.geogebra.org
[5] Web 2.0 calculator, see https://web2.0calc.com/widgets

of error they made. On the top of each STACK question, students could click on a drop-down accordion that explained how to use the input syntax (Fig. 2).

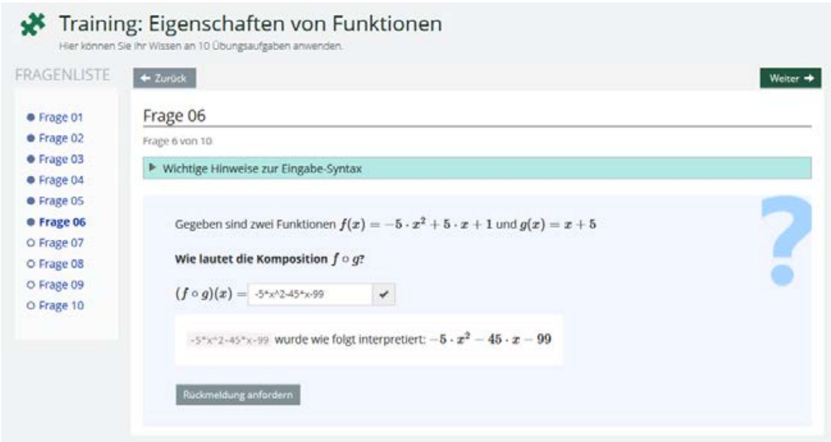

Fig. 1: Screenshot STACK item in course Functions, Training 1 ‘Properties of functions’

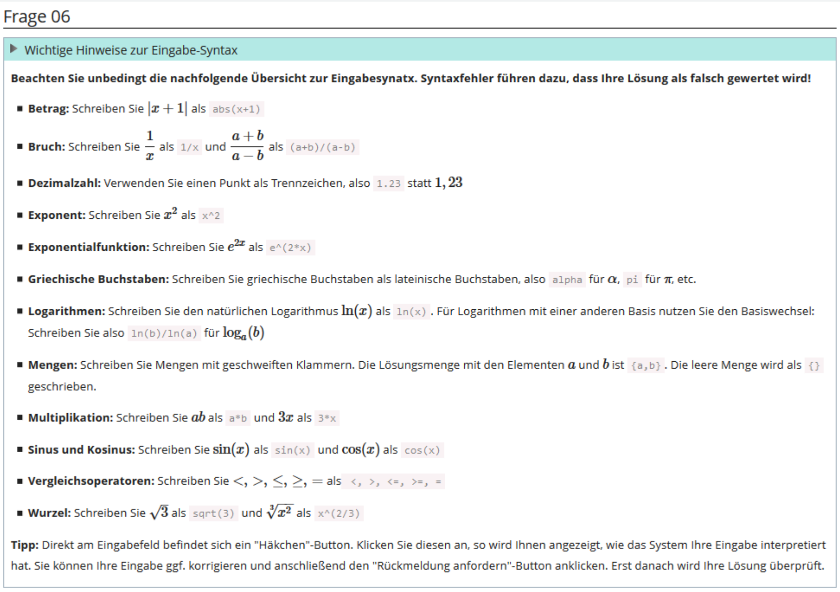

Fig. 2: Screenshot STACK item: ‘accordion’ with AsciiMath input hints

3 Sample

In 2019, 582 out of 748 first year students at DHBW Mannheim's technical faculty[6] participated in the study preparation in mathematics[7]. An evaluation survey that was administered to pre-course participants near the end of the course was answered by 126 students (answer rate = 22%). In this survey, four questions addressed the trainings and students' satisfaction with the STACK question type. For the evaluation of STACK item results, we analysed all attempts on a training / an item in this training. Tab. 2 shows how many students accessed the 13 trainings containing STACK items (note that some students accessed a training more than once).

	Training 1	Training 2	Training 3	Training 4
Arithmetic	262 (376)	163 (203)	125 (147)	
Equations	184 (234)	90 (113)	58 (66)	
Powers, roots, logarithms	115 (142)	71 (94)	60 (71)	
Functions	90 (107)	49 (55)	38 (40)	44 (49)

Tab. 2: Participants (attempts) per trainings

4 Results

4.1 Item analysis

One assumption had been that STACK questions would be answered correctly less often than ILIAS standard question types. We therefore analysed the results of the 13 trainings that used STACK as well as standard questions. Overall, 16,106 item attempts were analysed for 36 STACK items and 84 standard items. Of these attempts, 4,564 were related to STACK, 10,542 to standard items. All items scored either 1 (correct) or 0 (incorrect). 58% of all STACK items were answered correctly, whereas 63% of the standard items were answered correctly. It thus seems that STACK items were indeed answered correctly less often than standard items. This difference, however, was not significant (*ANOVA*: F (1, 118) = 2.26, p = .14).

It should be noted that 3,323 of the analysed item attempts (or 21%) were not answered at all (n/a). We therefore investigated if STACK items had been omitted more often by students than other items. Tab. 3 shows that this was not the case. On the contrary, standard items were skipped or left unanswered more often than STACK items, but again, this difference was insignificant. The main characteristic of an unanswered item was their

[6] Electrical Engineering, Mechanical Engineering, Mechatronics, Business Administration and Engineering, Chemical Engineering, Computer Science

[7] Minimum requirement for being defined as 'participant': submission of the diagnostic pre-test and the final test

position in the test: students often stopped answering the trainings after six or seven items, thus items near the end of a training were omitted more often than others. We then repeated the comparison between item types with n/a excluded. Now, the difference between STACK item scores and standard ILIAS item scores had increased, and this difference was significant, with F (1, 118) = 16.34, p <.01) (see Tab. 3).

	score		(n/a)	score (n/a excluded)	
	correct	wrong		correct	wrong
All items in trainings	61.2%	38.8%	20.6%	77.1%	22.9%
• standard ILIAS items	62.8%	37.2%	22.0%	80.5%	19.5%
• STACK items	58.1%	41.9%	18.0%	70.9%	29.1%

Tab. 3: Score (correct/wrong answers) standard 84 ILIAS items versus 36 STACK items

After having analysed the overall scores of STACK items we investigated the type of errors students made when answering STACK items. As the analysis of the data from the ILIAS database was very time consuming we were only able to perform this analysis for one training (see also limitations section). We chose Training 1 'Properties of functions' from the course Functions. It consists of seven standard and three STACK items that demand different forms of syntax input. Wrong answers were categorised into mathematical errors, syntax errors and non-answered items. Tab. 4 shows the number of errors that could be related to syntax and Tab. 5 shows the type of errors students made. It is striking that each of these errors could have been avoided if students had clicked on the validation button next to the input field (see Fig. 1), as STACK would automatically have alerted a syntax problem.

	correct answer	wrong answer	syntax error	n/a
Item 1	76	20		11
Item 2	55	33		19
Item 3 (STACK)	57	26	3	24
Item 4 (STACK)	76	4	0	27
Item 5	62	17		28
Item 6 (STACK)	42	31	8	34
Item 7	56	17		34
Item 8	51	19		37
Item 9	55	17		35
Item 10	59	11		37

Tab. 4: Number of syntax errors in STACK items. Example: Functions, Training 1

	cases
Wrong answers in 3 STACK items	61
Syntax errors	11
* for multiplication missing	7
X instead of x	2
– instead of -	1
x^2 instead of x^2	1

Tab. 5: Types of syntax errors in STACK items. Example: Functions, Training 1

4.2 Evaluation survey

Analysis of the evaluation survey gave an insight into the student perspective. The online survey had been answered by 126 pre-course participants. In this group, 97 stated that during the pre-course they had studied with trainings and STACK items. These students were asked to express their agreement with four Likert-scaled statements, from 5 (totally agree) to 1 (totally disagree), see Tab. 6.

As shown in Fig. 3, the majority of participants did not encounter severe problems answering STACK items. For example, 63% agreed to the statement Q2 "I did not experience any problems", and another 22% "somewhat agreed". So the overall results supported the view that the syntax issue was less relevant than expected and that STACK questions do not overburden pre-course participants. At the same time, nearly 20% did agree or totally agree to the statement Q4 'Input of mathematical syntax was difficult for me', indicating that indeed some of prospective technical students are not familiar with AsciiMath.

#	Item	n
Q1	I used trainings for practising.	92
Q2	I did not experience any problems.	97
Q3	Questions in trainings were displayed correctly.	96
Q4	Input of mathematical syntax was difficult for me.	81

Tab. 6: Evaluation survey (satisfaction with STACK items):
'How much do you agree with these statements regarding trainings?'

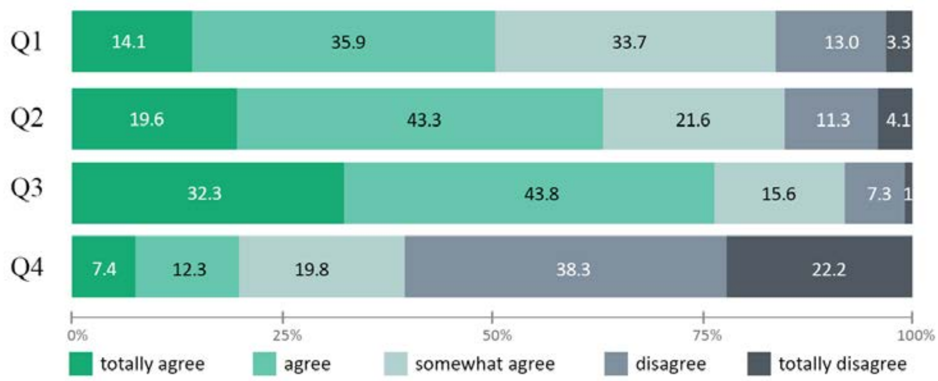

Fig. 3: Results evaluation survey (satisfaction with STACK items): agreement with Q1-Q4 in per cent.

5 Limitations

In this evaluation study not all quantitative data were available, or could be accessed in a reasonable period of time. Due to the organisation of test attempts, item attempts, and user IDs in the ILIAS database, it was extremely time consuming to evaluate the wrong answer patterns of STACK items. Thus our analysis of input errors was based on only one training. We hope that maybe the new version of STACK in ILIAS will allow for such analyses.

Furthermore, a comparison of answering times would certainly have been helpful for our evaluation, but ILIAS only documents the overall time of a test attempt but not the answering time per item.

Finally, the answer rate of the evaluation survey needs to be taken into account. While an answer rate of 22% may be acceptable to good in an e-learning environment (e.g. [CHT00]; [FY10]) the group of students that responded to the evaluation survey may not be representative of the student body regarding their mathematics or technical abilities.

6 Discussion

We investigated the suitability of STACK questions for a web-based pre-course in mathematics. For this evaluation study, 36 STACK questions were implemented into 13 trainings, in combination with 84 standard question types as provided by the LMS ILIAS (multiple-choice, single-choice, or numerical entry). Students answered these trainings in their self-study sessions. Support for AsciiMath input was provided via a drop-down button; there also was a help forum and a final evaluation survey. In this survey, feedback on STACK items was mainly positive. Against expectations, most students did not experience problems with this question type, and only a minority reported that they found the input of syntax demanding. While these students' problems should not be ignored, it also needs to be considered that all technical students will be confronted with mathematical tools in their first year at university, AsciiMath being the most basic of mathematical syntax.

Analyses of test results showed that STACK questions were not omitted or skipped more often than other items, thus students obviously did not avoid answering them. The rate of wrong answers, however, was higher for STACK items than for other item types, suggesting that either students made more syntax errors or that the authors more often chose the STACK question type for difficult mathematics problems. Effects of closed question types may have been influential, as well ([SJ17]).

An analysis of answer patterns of one training did not suggest that syntax errors dominated the wrong answers in STACK, most were related to the missing asterisk in multiplications. It was remarkable that nearly all of these errors could have been avoided if students had clicked on the validate button (Fig. 1) and corrected their input. We do not know why students did not use this feature; neither in the open answer section of the evaluation nor in forums or e-tutored courses the problem arose. Inexperienced e-learners have been found to be reluctant using help tools ([CE06]; [NPK07]) and might need support by human tutors, at least during the initial learning phase ([Az05]). An alternative explanation could be that pre-course participants overlooked the validation button as well as the accordion. We will not investigate this issue more deeply, as our 2020 version of the course will allow for direct validation, without the extra button.

Concluding, we consider STACK as perfectly suited for self-study environments but suggest that students should have access to a syntax help button. We would furthermore suggest to not include STACK items in diagnostic tests or test that are used to compare the mathematical knowledge of different learner groups, as indeed some pre-course participants struggled with the input. Standard question types like MC or numeric entry are sufficient for many problems, but by adding STACK questions the didactical variety of the optes trainings was much improved. We particularly appreciate the possibilities STACK gives for item feedback, so that each learner gets an appropriate and helpful feedback based on their input. In the final version of the optes pre-course STACK items will be added to all trainings, in particular for the extended curriculum and the topics vectors, linear algebra as well as continuous functions and limits.

7 Acknowledgements

Support for this publication was provided by the German Federal Ministry of Education and Research (BMBF) and its 'Quality pact for teaching' (ref. 01PL17012). Responsibility for the content published in this article, including any opinions expressed therein, rests exclusively with the authors. Test items and learning material are licensed under the CreativeCommons license BY-SA and can be provided via www.optes.de. The course

described in this paper can be accessed via https://studienstart.dhbw-mannheim.de.

Bibliography

[Az05] R. Azevedo, Using Hypermedia as a Metacognitive Tool for Enhancing Student Learning? The Role of Self-Regulated Learning, Educational Psychologist 40 (2005) 199–209.

[CE06] G. Clarebout, J. Elen, Tool use in computer-based learning environments: towards a research framework, Computers in Human Behavior 22 (2006) 389–411.

[CHT00] C. Cook, F. Heath, R.L. Thompson, A meta-analysis of response rates in web- or internet-based surveys, Educational and psychological measurement 60 (2000) 821–836.

[co14] cosh cooperation schule:hochschule, Mindestanforderungskatalog Mathematik (2.0) der Hochschulen Baden-Württembergs für ein Studium von WiMINT-Fächern, 2014.

[DHA18] K. Derr, R. Hübl, M.Z. Ahmed, Prior knowledge in mathematics and study success in engineering: Informational value of learner data collected from a web-based pre-course, European Journal of Engineering Education 10 (2018) 1–16.

[FY10] W. Fan, Z. Yan, Factors affecting response rates of the web survey: A systematic review, Computers in Human Behavior 26 (2010) 132–139.

[Ma99] M.E. Martinez, Cognition and the Question of Test Item Format, Educational Psychologist 34 (1999) 207–218.

[NPK07] S. Narciss, A. Proske, H. Koerndle, Promoting self-regulated learning in web-based learning environments, Computers in Human Behavior 23 (2007) 1126–1144.

[SJ17] C.J. Sangwin, I. Jones, Asymmetry in student achievement on multiple-choice and constructed-response items in reversible mathematics processes, Educational Studies in Mathematics 94 (2017) 205–222.

[Se13] SEFI mathematics working group, A Framework for Mathematics Curricula in Engineering Education, Brussels, 2013.

[Sw05] J. Sweller, Implications of cognitive load theory for multimedia learning, in: R.E. Mayer (Ed.), The Cambridge handbook of multimedia learning, Cambridge University Press, New York, 2005, pp. 19–30.

[WDH19] M. Weigel, K. Derr, R. Hübl, optes: Optimising Self-study With STACK, in: M. Sporring, C. Sangwin (Eds.), STACK Online Assessment. A collection of case studies, Edinburgh, 2019, pp. 35–40.