Ömer Genc, Ingmar Metzler (Hrsg.): In: Contributions to the International Meeting of the STACK Community 2021. TTK University of Applied Sciences: Tallinn, Estonia., Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2021 1

Effects of digital tasks on learning math for first-year engineering students

Ömer Genc[1], Ingmar Metzler[2], Ulrich Reif[3]

Abstract: This paper describes the conceptualization of digital tasks with STACK aiming at supporting tutors and students in entry-level math courses for engineering studies. The tasks are created within the project TU-WAS at the Technical University of Darmstadt (TU Darmstadt). Its agenda is to design, test, and assess digital tasks with potential effects on both tutors regarding their duties in exercise-units and students regarding their learning behavior as well as periods of self-study during the semester. Concerning statistical results are presented and a glance at the development of digital tasks is provided.

Keywords: computer aided assessment; mathematics for engineering; digital task design; feedback.

1 Introduction

Mathematical lectures for first year students in engineering subjects at German technical universities often have a large number of participants. In the winter semester 2020/2021 at TU Darmstadt there were 1620 students enrolled in the lectures "Math I for Civil Engineering (CE) / Mechanical Engineering (ME)". By most of these students, learning math is perceived as an obstacle, especially in their first two semesters. Difficulties in learning math among first-semester students are mostly characterized by heterogeneous and deficient knowledge of school-math [Ab14], which leads to a high dropout-rate from university and to lower grades in exams [He17]. It is shown that students mostly have problems with procrastination during the semester and that there is no regular learning and exercising [Sc13][Ge20a].
Due to the high number of students in entry level math courses for engineering, it is challenging for teachers and tutors to give an adequate and individual feedback to the participants. Given the fact that feedback lacks in quantity and quality [Fr16], it is no surprise that self-study and independent learning are not developed and encouraged at an adequate level. Feedback, in particular individual feedback, is crucial for supporting learning and increasing motivation [Ha07]. Using digital tasks to provide an individual feedback is the reason to focus on computer aided assessment.
The development of a computer aided assessment for supporting university courses leads to different research interests concerning the effects of digital STACK-based tasks on tutors regarding their work and on students regarding their learning behaviour. These research

[1] Technische Universität Darmstadt,genc@mathematik.tu-darmstadt.de
[2] Technische Universität Darmstadt,metzler@mathematik.tu-darmstadt.de
[3] Technische Universität Darmstadt,reif@mathematik.tu-darmstadt.de

doi:10.5281/zenodo.4916085

interests and hypotheses including the methodological procedure for the statistical evaluation and a brief example of an exercise are shown in the following sections.

2 Project TU-WAS

To support the students in learning and the tutors in working, the project "TU-WAS" was created at the TU Darmstadt [Ge20a]. The goal of the *TU-WAS*[4] project is to design, to test and to evaluate a collection of digital tasks for entry level math courses in engineering subjects.

2.1 Employment of STACK

The main goal for the project is to support students by initiating and strengthening phases of self-study and increasing the motivation by giving a well prepared and individualized feedback [Ha07]. Target groups are the courses Math I/II for ME, where the central topics are fundamentals (basic arithmetic, functions, set-theory), traditional and higher dimensional differentiation integration, power series and linear algebra.
The generated STACK-tasks are embedded in two different scenarios: as digital tasks for homework and as tasks in a digital learning environment where the students are able to use tasks for self-study in order to repeat specific topics from the lecture [Ge20b]. Tasks embedded in these two scenarios vary with respect to the amount of feedback given and the way the task is formulated. While designing and embedding tasks for a specific scenario, it is necessary to distinguish between different types of tasks (learning, performing and diagnosis) for deciding when to develop and to use which task for a specific scenario. Before creating digital tasks it was necessary to examine which kind of tasks were used in prior entry level math courses. It was shown that up to 75 % of previously used tasks can be implemented in STACK [Ge20b]. For creating *good* digital tasks including an elaborated feedback for supporting learning it is important to design a well-thought decision tree behind each task. In the next section, we present the structure of two decision trees.

2.2 A concrete application

In this section concrete exercises are presented to illustrate a possible employment of STACK within entry level maths courses at university and to explain the typical structure of associated decision trees.

Note that the latter have been reduced for compactness.

Figure 1 presents a sample exercise and depicts its associated decision tree. The latter is visually segmented to highlight components which appear frequently:

[4] Further information about the project, staff and the amount of resources: www.tu-was.jetzt

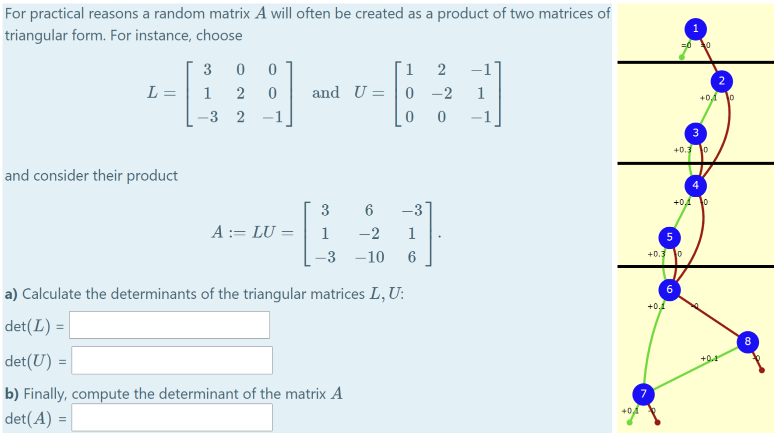

Fig. 1: Randomized STACK exercise utilizing an aftereffect check and its associated decision tree. The original exercise contains appended additional information concerning the LU decomposition.

1. Vertex 1 acts as a test for trivial answers in order to prevent students from dealing with the exercise carelessly. In the present case, it checks whether all entries are identical, to rule out the technically impossible cases $0, 0, 0$ and $1, 1, 1$.
2. The complex formed by vertices 2 and 3 represents a check for an individual answer. Vertex 2 checks whether an answer is algebraically correct, while vertex 3 checks whether the answer is adequately simplified. In case the expected answer possesses a canonical form, deviations to which are unacceptable, the *EqualComas* function may be employed. Otherwise, the number of operators appearing in the answer is compared with the number of operators remaining after Maxima internal simplification. If it does not exceed the latter plus an additional tolerance threshold, the answer counts as adequately simplified. For that purpose our team has developed a macro.
3. The complex formed by vertices 6,7, and 8 is an extension of the above and the most common structure for moderate feedback options. The purpose of vertices 6 and 7 is the same as of 2 and 3 above. However, in case the answer is not algebraically correct, Vertex 8 checks for aftereffects of previous mistakes and grants points in case the answer is correct based upon past mistakes.[5] Vertex 8 would also be used to check whether common, anticipated mistakes[6] have been made and not necessarily refer to Vertex 7! Compare the example below for details.

[5] The attentive reader might object that in the presented basic example vertex 8 is superfluous to check whether the product rule was correctly applied, as this might as well be checked directly in vertex 6. Since trivial answers were already ruled out, this is technically true. However, this incorporation is not always possible and exercises are designed to be extended or altered by collaborators to suit the purpose of the course they are meant to serve. For that reason, sticking to a sober scheme for decision trees is key to enable a third person to swiftly comprehend their structure and adjust an exercise.

[6] For instance, swapping basis matrices, expecting the determinant function to be additive or similar.

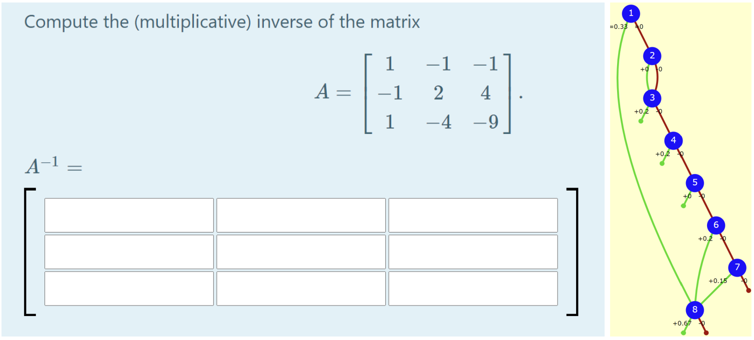

Fig. 2: Randomized STACK exercise for optimizing computation skills in preparation for examinations.

The tree in Figure 2 depicts an inflated version of Item 3 in the list above. Here, vertices 1 and 8 play the role 6 and 7 have played before. Vertex 2 solely serves the purpose of providing standard feedback for highlighting incorrect matrix entries and is not incorporated into Vertex 1 since highlighting is preferred *after* the individual feedback. Vertices 3 to 5 test for methodical mistakes providing hints tailored for eliminating underlying misconceptions. If there were no anticipated misconceptions, vertices 6 and 7 check if the majority of entries coincide with these of the teachers answer in order to prevent frustration due to slips. However, besides randomization and checking incorrect answers for aftereffects of previous mistakes or common misconceptions, a major advantage of STACK being CAS-based is that students inputs may be *qualitatively* tested. In fact, this allows posing much more open questions which demand deeper levels of understanding and challenge students to critically engage the course material. As the authors realize the value of this type of exercise, these are deployed frequently. However, none were suitable for illustrating the employment of STACK in a sufficiently compact and effective manner for these proceedings.

3 Effects of digital tasks on learning

3.1 Research interests and hypotheses

The embedments of digital tasks differ from each other. Therefore, different research interests prevail. Where digital tasks are used as home exercises, it is intended to figure out which possible effects STACK tasks have on work of tutors regarding their time of correction and the structure of their exercise units. For instance, a further examination can be conducted to identify, whether there is a relief in the tutors' work, so they can concentrate more time on giving feedback due to a decreased amount of printed home-exercises. In the other scenario the goal is to examine potential effects on learning strategies and especially on studying with tasks regularly during the semester. These research interests lead to the following four hypotheses built upon the potential of STACK concerning diagnosis and support [Ge20]:

1. STACK-based tasks (used in this setting) are able to initiate and intensify **phases of self-study** during the semester.
2. STACK-based tasks are able to increase the **activity on working with tasks**.
3. STACK-based tasks as digital homework are able to **relieve** tutors regarding their work on **correction of homework**.
4. STACK-based tasks are **appreciated** and seen as **helpful** by the students.

3.2 Methodological procedure

To answer these hypotheses above, two different questionnaires were created: one for students and one for tutors. In order to measure potential effects, these questionnaires were given to students and tutors from the course with digital tasks (*Math I for ME*, study group) and to two courses without digital tasks (*Math I for CE* and *Math I for Engineering at the TU Clausthal*[7], control groups) in the winter-semester 2020/2021. Both questionnaires consist of items with a five-level Likert-scale with total agreement or total disagreement to given statements [8]. For a longitudinal analysis the questionnaire for the students was presented twice: at the beginning and the end of the semester. In Figure 3 the timetable and the number of all students are shown. The first number shows the number of participants of the questionnaire, the second number shows all enrolled students, and the third number shows all active students based on *moodle*.

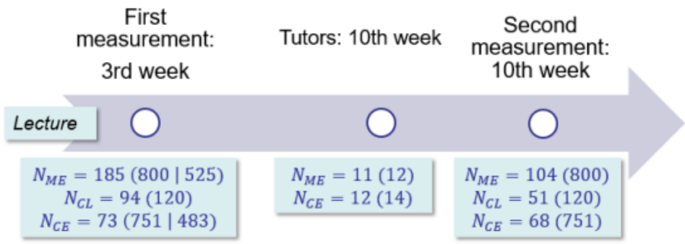

Fig. 3: Timetable of measurement

Learning at the university demands to control learning-procedures individually, which is often challenging for students [Vi90]. To initiate and intensify regular learning during the semester, it is necessary to examine, whether and which manifestation of cognitive and behavioral learning activities exist, and how they can be influenced by digital tasks. For this purpose an adapted version of the *LIST* questionnaire [Sc94] was used to examine different

[7] To have a higher number of participants for the questionnaire, the students from TU Clausthal were also asked to answer the questionnaire. Due to the same content of the lecture and the performance of the students, the cohorts can be compared.

[8] All items for both questionnaires can be found at: https://www.mathematik.tu-darmstadt.de/fb/personal/details/oemer_genc.de.jsp

learning-strategies. The questionnaire consists of 62 items, from which 28 items in different subscales were used from the *LIST* questionnaire. To measure further potential effects, items concerning motivation (8 items), knowledge on specific mathematical topics (12 items) and general acceptance (8 items) were added to the questionnaire. The questionnaire for tutors consists of 6 items concerning the time spent on correction and the amount of given feedback. A confirmatory factor analysis shows acceptable fits based on commonly used criteria $\left(\frac{\chi^2}{\text{df}} < 2.5; \text{CFI} > 0.9; \text{TLI} > 0.8; \text{RMSEA} < 0.08\right)$. The comparison to prior used questionnaires for measuring learning strategies using *LIST* or adapted *LIST* versions ensures reliable subscales used in the questionnaire [Sc94][Bo05][Kl18]. A brief insight into the results of the study are given in section 3.3 below.

3.3 Effects on learning-behavior and self-study

To answer the first two hypotheses by using the subscales in the questionnaire, new scales based on the reliable subscales were created and put together. For each of those subscales the inner consistency was ensured ($\alpha_{\text{Cronbach}}, \omega_{\text{McDonald}} > 0.7$). A Wilcoxon-Mann-Whitney-Test first within and then between the study and the control group was used to examine potential (significant) differences of changes of the middle rank in the pre- and posttests (i.e. the two different measuring points) based on mean-values from the items of each subscale in favor of the study-group. In Table 1 exemplar results are shown with $\text{N}_{\text{SG}} = 101$

	U	Z	Significance	Pearson r
H1: Cognitive learning-strategies	3675.000	-4.582	0.000	0.312
H1: Meta-cognitive learning strategies	4386.500	-3.025	0.002	0.206
H1: Internal ressource-related: time	4825.000	-1.954	0.044	0.133
H1: Motivation	4605.500	-2.536	0.011	0.173
H2: External ressource-related: tasks	4760.000	-4.987	0.027	0.340
H2: External ressource-related: fellow students	5589.500	-0.009	0.993	0.000
H2: Internal ressource-related: effort	4964.500	-1.654	0.098	0.113

Tab. 1: Exemplar results of Wilcoxon-Mann-Whitney-Tests for hypothesis 1 and 2

and $\text{N}_{\text{CG}} = 114$. Based on the results, Hypothesis 1 can be verified: there is an significant increase of usage of cognitive, meta-cognitive and resource-related learning strategies and motivational aspects between the study- and control-group in favor of the study-group. Hypothesis 2 can be also verified partly, considering that there is a significant increase of usage of resource-related learning strategies (working with tasks) between study- and control-group in favor of the study-group. In addition, a t-test and a Welch-test both show a significantly ($p < .001$) higher rate of activity on STACK-tasks and submitted STACK-tasks for homework in comparison to ordinary tasks within the study-group. For all other learning strategies and potential changes of mathematical knowledge there is no significant increase measured.

3.4 Effects on tutors and correction time

To answer Hypothesis 3 a second questionnaire was given to tutors of courses of both groups concerning time of correction and the quality and quantity of given feedback while correcting homework. In Table 2 the results are shown with $N_{SG} = 11$ and $N_{CG} = 12$. It can

	U	Z	Significance	Pearson r
H3: Correction time 1	15.500	-3.216	0.001	0.671
H3: Correction time 2	20.000	-3.000	0.001	0.626
H3: Correction: Quality and quantity 1	43.500	-1.462	0.144	0.305
H3: Correction: Quality and quantity 2	64.000	-0.064	0.0949	0.013
H3: Correction: Quality and quantity 3	54.500	-0.730	0.466	0.152
H3: Correction: Quality and quantity 4	55.500	-0.657	0.511	0.137

Tab. 2: Results of Wilcoxon-Mann-Whitney-Tests for hypothesis 3

be shown that the amount of time needed for the correction is significantly lower for tutors in the study group than in the control group, although there is no significant difference concerning the quality and quantity of correction between two groups. That means, while tutors of both groups gave a comparable amount of feedback, the tutors of the study group needed less time for correcting and giving feedback.

3.5 Acceptance in digital tasks

While the prior hypotheses examine potential effects on learning and teaching, the fourth hypothesis deals with the students' acceptance of STACK tasks. For answering this hypothesis, six items were added to the questionnaire considering the appreciation and the usefulness of the digital tasks. A t-test with $N_{SG} = 101$ was used to evaluate the results. It can be shown that the students significantly appreciate digital tasks and that they would like to use STACK tasks for learning and preparation in other courses. In two open-answering questions, students were asked to describe why they might appreciate the concept of digital tasks, and what should be done better in the future. Here, the most popular answers were that digital tasks delivered a transparent and immediate feedback, and enabled training possibilities for certain topics. Suggestions for improvement were that students desire more elaborated feedback regarding delivering sample solutions, or explicit hints on how to solve problems and not only to determine what they probably did wrong.

4 Conclusion and future work

The next step for the project is to repeat the whole setting with digital tasks for math II courses in the summer semester of 2021, and to gather further data with the use of the questionnaires. The questionnaire itself will be improved. That means that certain items with low validity and correlation will be replaced or removed in order to achieve a better validity and selectivity in the different subscales within the questionnaire for answering the hypotheses. After improving all used tasks and questionnaires, and while examining the results after the first two semesters, the whole setting will be repeated for the same courses in the winter semester of 2012/2022, and in the summer semester of 2022 to obtain a reasonable amount of data for examining potential effects of digital tasks on the learning strategies being investigated.

Bibliography

[Ab14] Abel, H., Weber B.: 28 Jahre Esslinger Modell – Studienanfänger und Mathematik. In (Biehler, R. et al., ed.): Mathematische Vor- und Brückenkurse, Konzepte und Studien zur Hochschuldidaktik und Lehrerbildung in Mathematik, pp. 9–19. Wiesbaden: Springer Spektrum, 2014.

[Bo05] Boerner, S., Seeber G.-Keller H. & Beinborn P.: Lernstrategien und Lernerfolg im Studium. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 37:17–26, 2005.

[Fr16] Freyn, W., Weiß C.: Neue Maßnahmen für eine verbesserte Schulung und Betreuung von Übungsleitern. In (Hoppenbrock, A. et al., ed.): Lehren und Lernen von Mathematik in der Studieneingangsphase, Konzepte und Studien zur Hochschuldidaktik und Lehrerbildung Mathematik. Wiesbaden: Springer, 2016.

[Ge20a] Genc, Ö.: Projekt TUWAS: Einsatzszenarien und mögliche Effekte STACK-basierter Mathematikaufgaben im Ingenieurstudium. In: Beiträge zum Mathematikunterricht 2020. 2020.

[Ge20b] Genc, Ö.: STACK for mathematics in engineering: concepts and effects for teachers and students. In: Contributions to the 3rd International STACK Conference 2020. 2020.

[Ha07] Hattie, J., Timperley H.: The Power of Feedback. Review of Educational Research, 77(1):81–112, 2007.

[He17] Heublein, U., Schmelzer R.: Die Entwicklung der Studienabbruchquoten an den deutschen Hochschulen. Berechnungen auf Basis des Absolventenjahrgangs 2016. DZHW Projektbericht, 2017.

[Kl18] Klingsieck, Katrin B.: Kurz und knapp - die Kurzskala des Fragebogens "Lernstrategien im Studium". Zeitschrift für Pädagogische Psychologie, 32(4):249–259, 2018.

[Sc94] Schiefele, U., Wild K.-P.: Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. Zeitschrift für Differentielle und Diagnostische Psychologie, 15(4):185–200, 1994.

[Sc13] Schulmeister, R.: Auf der Suche nach Determinanten des Studienerfolgs. In (Pilniok, J. Brockmann/A., ed.): Studieneingangsphase in der Rechtswissenschaft, pp. 72–205. Nomos: Baden-Baden, 2013.

[Vi90] Viehbahn, P.: Psychologie studentischen Lernens. Weinheim: Deutscher Studien Verlag, 1990.