

Multiple Choice Tests: More than a Time Saver for Teachers

<http://dx.doi.org/10.3991/ijep.v5i3.4376>

Wolfgang Fellin and Gertraud Medicus
University of Innsbruck, Innsbruck, Austria

Abstract— Multiple choice tests (MC tests) are usually used as a tool for assessing factual knowledge in courses with many students. This article proves that MC tests bear a much greater potential for education in general and for the training of engineering students in particular. To justify this claim, two different MC tests were developed, consisting of a large pool of questions that assessed a wide range of cognitive competences needed in geotechnical engineering. The novel feature of the tests can be seen in the immediate comments given to right or wrong answers in order to understand the correct solution. One test was used as a geotechnical quiz, the other one as an online self-assessment tool in engineering courses attended by bachelor students. The expected effects of the MC tests were tested against a number of hypotheses based on feedback data from students' questionnaires and interviews. The major findings were the following: The MC tests were not used throughout the courses, they were used shortly before the written exams. The tests were seen as a valuable tool for preparing for the exam due to the crucial feedback feature of the test. The tests helped to deepen the understanding of theoretical concepts, increased students' interest in geotechnical engineering and slightly raised the performance level of bachelor students in the written exams. The answers to the interviews also revealed that students were aware of the danger of using sample solutions for the exam at the cost of not developing a deeper understanding for theoretical concepts. The authors are convinced that the deployment of MC tests as a tool for blended learning and self-assessment will have long-term training effects and therefore justify the amount of work required for constructing such tests.

Index Terms—blended learning, learning management system, learning on demand, self assessment

I. INTRODUCTION

We observed that civil engineering students tend to distinguish between theory and practice. The foundations for developing geotechnical models and the manipulation of equations are called theory. When students can put in some numbers they call it practice. Our students highly prefer practice. This – in our opinion – artificial splitting is somehow amplified by the traditional teaching system in Austria: lectures, in which theory is presented by professors, and exercises, in which students perform calculations.

We wanted to bridge this gap in a playful way. For this purpose, we designed a multiple choice self-test, available online via the learning management system OLAT (Online Learning And Training) [1], called “geotechnical quiz”. To include the quiz into the written exam, we modified this test to a mixture of more complex calculation examples and a multiple choice part. The multiple choice (MC) part comprises questions from the quiz, additional basic

modeling issues and simple calculations. As the students performed rather poorly in the multiple choice part, we developed a new online self-assessment and training tool: a MC test as a kind of pretesting strategy, which is known to boost the performance in subsequent assessments [2].

Depending on the content of the tasks [3], MC tests can be used to assess higher cognitive skills rather than focusing on knowledge recall [4]. Our MC test corroborates this assertion: Only a few of the MC questions tackle concrete facts, the majority of questions require short analytical or numerical calculations, i.e. they represent simplified versions of more complex engineering calculations.

MC tests may have negative effects, e.g. the so-called negative suggestibility effect, i.e. students will sometimes come to believe that distractors are correct, and therefore acquire false knowledge [5]. This occurs when the distractors are chosen erroneously, reading the distractors on their own does not have a negative effect [5]. Moreover, the negative suggestibility effect can be eliminated by means of feedback on the test [6], which is the case in both of our tools. The detailed feedback provided by our geotechnical quiz also plays an essential role in the learning process [7], whereas simply labelling answers as true or false does not improve performance on the consecutive tests. The positive effects of MC tests generally outweigh the negative ones [8], [9], [2] and MC tests are regarded as part of high quality university teaching [10].

Both online MC test tools are used in addition to the regular face-to-face courses in which examples of traditional engineering calculations are trained, in the sense of blended learning. The course is offered in the third semester of the bachelor program in civil engineering science, and covers the following basic geotechnical topics: physical parameters of soils, classification of soil, seepage flow, seepage force, effective and total stress, stress distribution in the ground, calculation of settlement, consolidation and creep, shear strength, earth pressure. The training examples are selected from a pool of old examination examples (see online [11]). Students calculate the examples in between the course meetings based on the lecture of the professor, lecture notes and engineering standards (i.e. Euro code and Austrian national standards). The lecture notes are based on [12] and are continuously corrected and further developed, e.g. changes in the standards are implemented. Since 2003, the course program, attended by up to 36 students, is structured as follows. A

short “warm up” with small portable geotechnical experiments (compare [13], [14], [15]), quizzes (in the format of “Who Wants to Be a Millionaire?”) and case studies of damages (students act as geotechnical experts in court). After this warm up, small groups of 4 students discuss their calculated examples. The students are advised to compare not only the result of each example, but also the applied calculation method. The group should then decide which result and calculation method is correct for each example. The teacher gives support if the group is unable to solve the questions appearing through group discussion. To stimulate the discussion, the teacher actively poses additional questions related to the calculation examples and the underlying theories. The students also have to solve or answer up to 6 additional small calculations or questions, which are handed out at the beginning of the discussion period. Randomly selected students have to present their calculation method and results on the blackboard. The course events end with answering the additional questions/examples by randomly selected students or, if necessary, by the teacher. The structure of the lecture is therefore in line with the findings of Alam and Jackson [16], who showed that hands-on experiences and demonstrations and face-to-face feedback increase the motivation of students to attend lectures.

II. GEOTECHNICAL QUIZ

The implementation of the geotechnical quiz was motivated by the following hypotheses:

- 1) Despite not being mandatory, the quiz will be used throughout the course.
- 2) Assumed reasons for the usage are: learning on demand, joyful learning, and a good preparation for the written exam.
- 3) Interest in geotechnical engineering will increase. (Note that, geotechnical engineering courses count only for 10 ECTS credits out of 180 ECTS credits of the whole bachelor program.)
- 4) The quiz facilitates the application of theories and improves their understanding, which is essential in complex geotechnical calculations.
- 5) The test serves as a good preparation for the written exam.

We¹ had to implement the self-test in the learning management system OLAT [1], which is the standard e-learning platform used at the University of Innsbruck. We constructed about 100 multiple choice, single choice and cloze test questions on basic modeling issues and short calculations [17].

A. Implementation

An example of a quiz question and its implementation is given in Fig. 1. Detailed feedback is provided for answers that may be chosen by the students, see e.g.

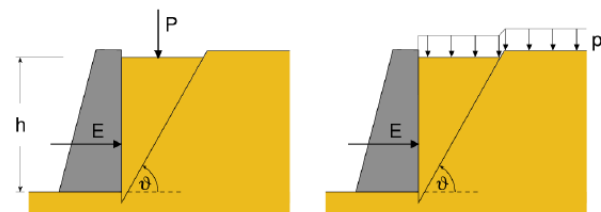
¹The quiz is based on an idea of the first author and has been worked out by the second author as an action research [23] project during her professional training “Zertifikat Lehrkompetenz” (teaching skills certificate: <http://www.uibk.ac.at/personalentwicklung/lehrkompetenz/zertifikat.html>) at the University Innsbruck.

Fig. 2, which shows the feedback to the multiple choice question in Fig. 1. Feedback is an important feature of online tests (e.g. [18], [19]) and should be given immediately to the test answer [20]. Questions, answers and feedback are visualized by means of figures, pictures and embedded movies. Most figures were taken from [21] and the majority had to be slightly adjusted for the quiz.

Influence of a surcharge

How does a surcharge P (kN/m) or p (kN/m²) on the ground surface affect the earth pressure?

Figure: Determination of the active earth pressure E when the surface is loaded.



- The active earth pressure increases.
- The passive earth pressure increases.
- The passive earth pressure decreases.
- The active earth pressure decreases.
- A surcharge has **no influence** on the earth pressure.

Figure 1. Example of a question in the OLAT geotechnical quiz.

The quiz was first implemented in the winter semester 2011. The motivation of the students to use the new tool was rather low until we informed them that some of the questions would also appear in the written exam.

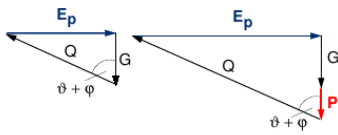
B. Evaluation

An evaluation conducted after the examinations in 2012 obtained a generally very positive feedback, e.g. 80% of the users found that doing the test was at least partly joyful and 90% were under the impression that the test helped them to prepare for their exams. A reevaluation in 2014 generally confirmed the result of 2012, but the level of satisfaction was slightly lower than in 2012. Additional interviews in 2014 with 10 students (see interviews on online multiple choice test below) revealed a very positive acceptance of the quiz: students liked to use the test and found it very helpful for learning and understanding the underlying theories.

1) *Paper and pencil poll:* A paper and pencil poll using a questionnaire was carried out in the follow-up lecture. 90 forms were collected in 2012 and 107 in 2014. The questionnaire posed 8 closed and three open questions. The answers to the closed questions are summarized in Tables I - II and Fig. 3 and 5. The majority

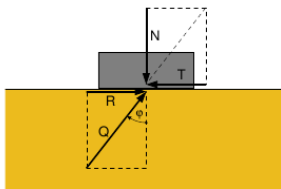
☒ The passive earth pressure decreases.

Wrong! Please compare figures below:



Tip:

To understand the direction of the force Q, compare:



To better understand the difference between (a) active and (b) passive (earth)pressure, compare:

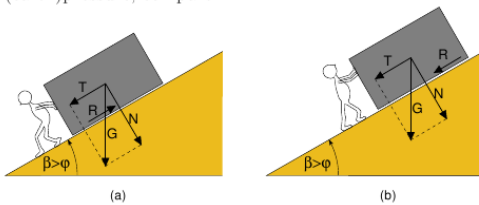


Figure 2. Example of a feedback to a wrong choice.

(75% of the students in 2012, 57% in 2014) answered the open questions.

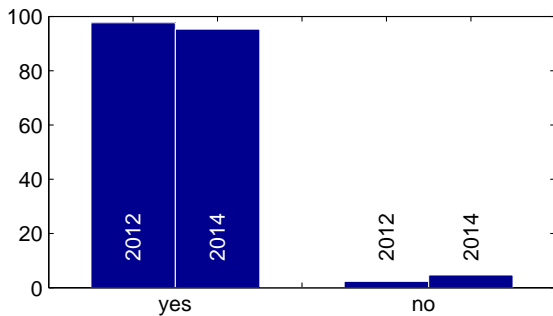


Figure 3. Percentage rate regarding the question: Have you used the quiz for exam preparation?

Almost every student used the quiz for exam preparation, see Fig. 3 and Table I. In Fig. 4 the access statistic² of the geotechnical quiz is shown. Continuous use during the semester was scarce. There was an initial period of interest during December after the presentation of the new tool, but students did not continue using the quiz. We therefore announced in January that some of the questions will appear in the final exam, which slightly increased the number of accesses and finally resulted in the peaks of usage shortly before the exams, see also Table I.

²Note that the data in Fig. 4 only give information about the frequency of access to the quiz and not about the number of questions answered or the usage duration.

The vast majority (89.9% in 2012, 74.3% respectively in 2014) felt better prepared for the exam by using the quiz, see Fig. 5.

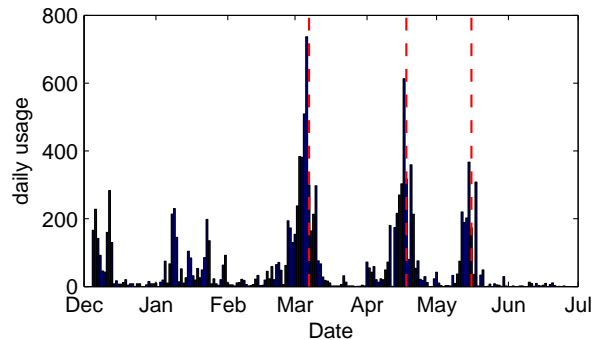


Figure 4. Daily usage of the online geotechnical quiz from the event log of the software (2012). Dashed red lines mark the dates of the written exams.

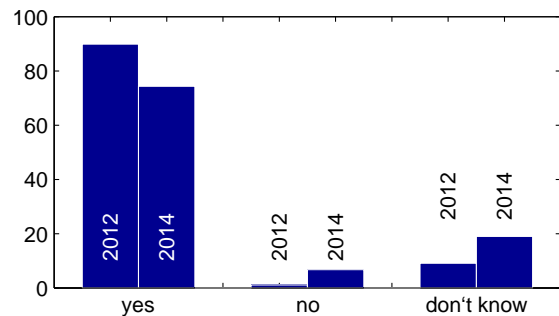


Figure 5. Percentage rate regarding the question: Are you under the impression that you were better prepared for the exam by using the quiz?

Around 90% of the students agreed or partly agreed that the quiz facilitated the use of theories and improved their understanding, which was required for the calculations. Around 10% disagreed or could not estimate the influence, see Table II. 20% (2014) to 30% (2012) agreed that it was fun to do the quiz.

The question “What did you like?” was answered by 58% of the students in 2012 and by 49% in 2014. The explanations/feedback function was mentioned to be helpful, especially for the wrong answers. Furthermore, they liked the clear illustrations, the easy handling and the flexible and independent usage.

The question “What did you dislike?” was answered by 36% of the students in 2012 and by 30% in 2014. They mentioned that the questions were too easy (compared with the exam questions). Some disliked the handling of the software for calculations.

The question “What should be improved?” was answered by 39% of the students in 2012 and by 20% in 2014. Some wanted an increase in difficulty of the questions and some asked for further extensions of the test.

2) Interviews: The interviewed students required only minor further enhancement of the test:

PAPER
MULTIPLE CHOICE TESTS: MORE THAN A TIME SAVER FOR TEACHERS

Table I
FREQUENCY OF THE USE OF THE QUIZ.

	year	often (1)	hardly (2)	never (3)	don't know	mean	number of valid answers
How often have you used the quiz? during the semester	2012	12.6	32.8	49.5	5.0	2.39	89
	2014	9.0	37.6	50.5	3.4	2.42	107
How often have you used the quiz? shortly before the exam	2012	79.4	16.7	2.3	1.6	1.22	89
	2014	80.2	13.7	3.3	2.9	1.21	104

Table II
RATE OF AGREEMENT TO EACH STATEMENT: NUMBER OF ANSWERS RELATIVE TO THE VALID ANSWERS.

	year	agree (1)	partly agree (2)	disagree (3)	don't know	mean	number of valid answers
The quiz facilitates the use of theories which are required for the calculations.	2012	47.7	43.0	4.7	4.7	1.55	90
	2014	32.4	53.3	12.4	1.9	1.80	105
The quiz improves my understanding for the calculations presented in the lecture.	2012	54.7	34.9	3.5	7.0	1.45	90
	2014	29.5	59.0	7.6	3.8	1.77	105
Working with the quiz provided a good preparation for the subsequent course.	2012	46.5	39.5	3.5	1.05	1.52	90
	2014	39.0	35.2	13.3	12.4	1.71	105
The quiz increased my interest in geotechnical engineering.	2012	19.8	40.7	26.7	12.8	2.08	90
	2014	11.4	41.0	35.2	12.4	2.27	105
It was fun to do the quiz.	2012	27.9	52.3	12.8	7.0	1.84	90
	2014	18.3	49.0	22.1	10.6	2.04	104

For the computational tasks the students asked for more hints and feedback. Moreover, the procedure for computational tasks should be facilitated. Another criticism for the computational tasks was that the software requires exact numbers and accumulated rounding errors cause differences in the solutions. It is thus desirable that the software could check a number range instead of only one exact number.

Some of the students asked for more test questions, which should be a little more challenging, i.e. comparable to the online multiple choice test.

3) *Further impressions:* Some workaround solutions were attempted by students, like the request "Can we get a pdf-version of all questions of the quiz?". Students used the online test very often to find out correct solutions by accident and collected screen shots of them. Such lists could probably be used as a data basis for right and wrong answers, which shows the intention to learn answers and solutions by heart, without understanding them.

It was very interesting to notice that the students used some questions of the quiz and their answers in discussions of some related geotechnical problems in the follow-up course, long after the examinations.

4) *Interpretation and consequences:* Almost every student used the test. However, the use of the test shortly before the exam was much more intensive than the continuous use. Thus, our first hypothesis is only partly confirmed. For the majority of students the self-test serves as an aid for understanding the calculations based on theories (hypothesis 4) as well as for the subsequent course. Almost all students thought they would be better

prepared for the exam by using the quiz (hypotheses 2 and 5). Every third student had fun doing the test (hypothesis 2). The aim to increase the interest in geotechnical engineering by using the test has only been partially achieved (hypothesis 3). Just over half of the students agreed at least partly with the statement that the quiz increased their interest in geotechnical engineering (18% agreed, 35% partly agreed).

Interestingly, the satisfaction slightly decreased from 2012 to 2014. The reason for that could be the following: In 2011/2012 students used the quiz for the first time. We asked them to improve the test on the basis of their own ideas for questions or by finding (spelling) mistakes. Students who participated in the improvement were given prizes in a raffle. The personal involvement could be one of the reasons why they were more satisfied in 2012, even though the test contained more errors than in 2014. Another explanation could be a ceasing novelty effect for those students who took the course a the second time.

Concerning the evaluation in Table II a four- or five-step scale would allow a more precise assessment, similar to the one for the online multiple choice test in Table III. For the reevaluation in 2014, the same questionnaire was used in order to achieve a better comparability.

As part of an e-learning project, the OLAT test was transferred in the ONYX software³. The handling is more intuitive and easier than the one of OLAT. However, it is not possible to include videos and figures as feedback which are indispensable for the geotechnical quiz. We

³www.onyx-editor.de

therefore still use the OLAT test, until the ONYX software meets the desired requirements.

In the future, there will be improvements concerning the handling of computational tasks.

With the included explanations and feedback, the geotechnical quiz serves as an appropriate online learning tool with a fair number of questions. Concerning the students' request for more complex questions, we have to inform the students that the geotechnical quiz is a learning tool and that it is not the only appropriate tool for exam preparation. The geotechnical quiz is intended to serve as a supplement to the online multiple choice test (see Section III), which contains more challenging questions.

III. ONLINE MULTIPLE CHOICE TEST (MCT)

As stated before, we promised to use questions of the geotechnical quiz for the written exams to enhance the motivation of the students to use the quiz. So as to include the questions of the geotechnical quiz into the written exam, we changed this test into a mixture of longer calculation examples and a multiple choice part. The multiple choice part comprises questions from the quiz, as well as additional basic modeling issues and short calculations. Characteristic tests can be found online on the homepage of our division [11]. As the students performed rather poorly in the multiple choice part, we⁴ developed a new online-training tool. This test should exactly simulate the multiple choice part of the written exam. The online layout is therefore exactly the same as in the written test, e.g. Fig. 6.

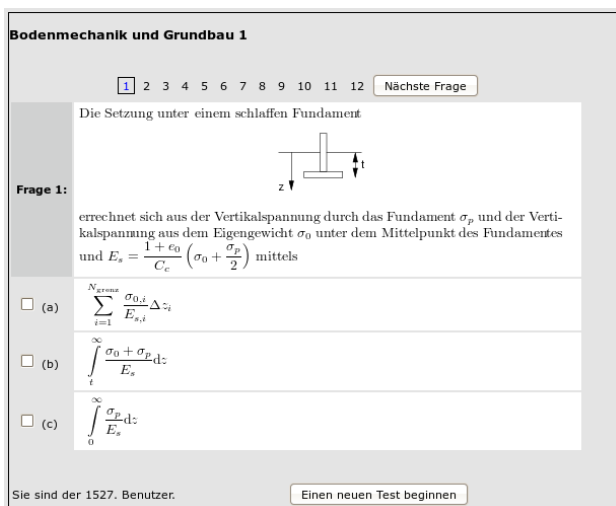


Figure 6. Example of an online multiple choice question.

The online MC test was introduced based on the following hypotheses:

- 1) The test provides a good self-assessment.
- 2) The test enhances skills for more complex geotechnical calculations.

⁴A software, which was used by Tobias Hell for the courses Analysis 1 and 2 at the Institute of Mathematics of the University of Innsbruck, was further developed and extended in an e-learning project led by Tobias Hell and the first author [22]. The programming was performed by Gregor Staggl.

A. Software and Features

A multiple choice test software program for Mathematics was further developed and extended to meet geotechnical needs [22].⁵ The software (HTML, PHP, AJAX) randomly chooses a user-defined amount of questions from a large pool (MySQL data base, entries in \LaTeX format). It can produce online tests (https://webapp.uibk.ac.at/geotechnik/mc_ue_bmgbl/) as well as printed versions for the real exam.

The main improvement of the new version is the introduction of larger pools of possible answers to each question, of which only a user-defined part is randomly chosen by the program and displayed in each realization of a test. This should avoid the tendency to learn questions and related correct answers by heart, as students very likely get different possible answers for the same question in successive tests. Therefore, they have to go through the theory or the short calculation again to solve the question correctly. Additionally, the hidden pool makes working out comprehensive standard solutions for students in the following years much harder. Each question can be labeled to appear either in the online test, in the written exam or in both tests. Questions appearing solely in the printed tests can be retained for the real exams. The same labeling is possible for each answer. Too many distractors have a negative effect on learning [5]. The number of answers in the test can be set by the teacher. We advise 3 or 4 answers in the test and a pool of 8 to 12 possible answers for each question. The teacher can also set the minimum and maximum number of correct answers that should be displayed.

We also offered the possibility to include figures in the questions and answers. This feature is essential to cover geotechnical issues, which was also the case in the above presented geotechnical quiz.

A new system of categories was introduced. Each question can be part of one user-defined category, and the test will be composed of a question from categories set by the user, so that a test covers the whole content of the lecture. This feature can be used to replace some categories in the written test as the topics are covered by the additional longer calculation examples.

As it is inherently difficult for teachers to guess the level of difficulty of specific questions for students, we implemented a counting system in the online test which counts the number of correct and wrong attempts for each question appearing in any realization of the online tests. This information gives a hint about the level of difficulty of each question. Obviously, one has to be aware of any bias, like the time the question exists in the test (which yields higher positive rates) or some thoughtless trials. However, with or without this information the teacher is supposed to choose a level of difficulty for

⁵It is worth to mention that a similar software TCexam (<http://www.tcexam.org/>) has been used since mid of 2014 at University Innsbruck for online exams. However, as our project had already started in 2012, we had to fit in the computer systems provided at that time, e.g. a special format for automatically scoring a paper and pencil version of the test.

each question from three difficulty categories (simple, intermediate, difficult) and to set the number of questions from each category which should appear in the tests. This should reduce the different difficulty levels of successive test.

In the first version we decided to stay close to the real test and display only the overall amount of correct questions evaluating the test. No further feedback was provided.

The preparation of the tool took about 280 working hours for the software improvement plus approximately 20 hours to import the 54 questions from the existing written exams and to generate additional answers for the pool of answers.

B. Evaluation

The evaluation was performed subsequently to the last written exam. Students had to pass one written exam. Three possible dates were offered, so that students could choose a convenient date, and had the possibility to repeat the exam once, in case of failing their first attempt.

1) *Paper and pencil poll:* A paper and pencil poll using a questionnaire was carried out in the follow-up lecture. 105 forms were collected. One form was excluded because the student claimed to have performed the test 840 times, which is very unlikely. In fact, the event log of the software recorded 1531 usages between 1st of January to 1st of June. The sum of the declared estimated usages was 2081 and 1241 with and without the high claim, respectively.

Students were asked to rate their degree of agreement to 9 statements, as displayed in Table III. Forms in which the number of performed online MCT is zero or not given are not considered in evaluating all statements. We asked students about their performance in the written exam. Forms in which students either claimed not to have participated in the written exam or this question was left unanswered were excluded from the evaluation of the statement "The online multiple choice test (MCT) served as good preparation for the written exam" (row 8 of Table III).

The reply to each further *open question* is summarised below. The answers to the question "Did you use worked-out standard solutions for the MCT?" were marked valid 83 times. Forms in which this item was not filled or the number of performed online multiple choice test is zero or not given were not considered in evaluating this question. The valid answers are: 6.0% always, 10.8% frequently, 25.3% seldom and 57.9% never.

The answers to the question "How often did you perform the multiple choice test?" were marked valid 93 times. The result is shown in Fig. 7: The majority (51,5%) used the test less than 10 times, only 2.2% more than 30 times. The mean usage was 13 times.

The question "Do you want to tell us which grade you achieved in the written exam?" was answered 71 times by revealing the grade, 8 students chose the optional answer "No, I did not take the written exam" and 22 chose the optional answer "No, I do not want to tell my grade." The

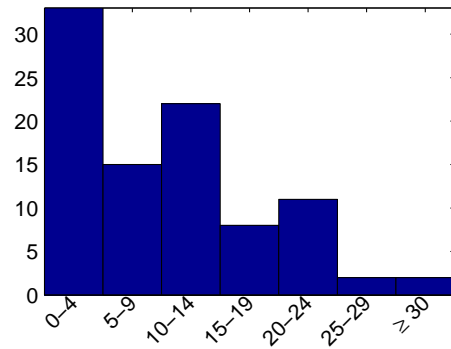


Figure 7. Declared number of attempts of the online MCT

disclosed grades are not correlated with the given number of test attempts, see Fig. 8.

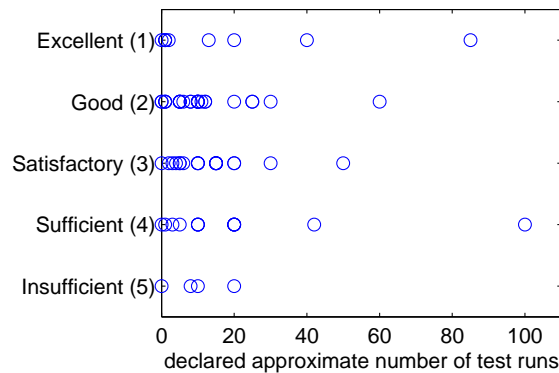


Figure 8. Correlation between achieved grade and declared number of MCT attempts.

The answers to the question "In case of redoing the written exam, did you use the MCP test ..." were marked 42 times: 28.6% more often, 23.8% equally often, 19.0% less often, 28.6% no longer. This is in agreement with the records of the event log of the software, see Fig. 9. With the second examination date approaching, the number of usages was much higher than before the first exam.

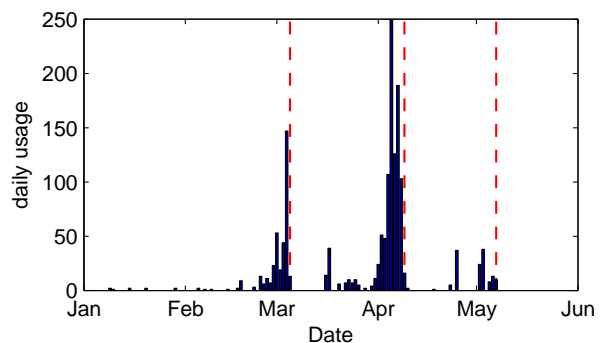


Figure 9. Daily usage of the online MCT from the event log of the software. Dashed red lines mark the dates of the written exams.

The question "What did you like?" was answered 22 times: 6 stating that the multiple choice test was a good preparation for the written exam, 5 appreciated

PAPER
MULTIPLE CHOICE TESTS: MORE THAN A TIME SAVER FOR TEACHERS

Table III
RATE OF AGREEMENT TO EACH STATEMENT: NUMBER OF ANSWERS RELATIVE TO THE VALID ANSWERS.

	agree (1)	somehow agree (2)	(3)	somehow disagree (4)	disagree (5)	mean	number of valid answers
I have learned a lot by working through the online multiple choice test (MCT).	2.0	29.3	26.3	29.3	13.1	3.2	99
The MCT was too easy.	0.0	0.0	19.0	40.0	41.0	4.2	100
The MCT improved my understanding of basic geotechnical theory.	3.1	30.9	23.7	26.8	15.5	3.2	97
The MCT makes solving longer calculation easier.	2.0	16.2	21.2	35.3	25.3	3.7	99
Working through the MCT was entertaining.	0.0	12.5	41.6	29.2	16.7	3.5	96
The MCT is a superfluous tool.	2.0	11.9	20.8	21.8	43.5	3.9	101
Working through the MCT was exhausting.	10.3	44.3	32.0	11.3	2.1	2.5	97
The MCT served as a good preparation for the written exam.	13.1	33.4	28.3	12.1	13.1	2.8	99
I would recommend the MCT to other students.	15.0	31.0	24.0	15.0	15.0	2.8	100

the existence and/or the principal setup of the MCT, 4 stating that the MCT was as challenging as the written exam, 4 acknowledged the plurality and the level of the questions, 1 observed that the answers were changing, 1 acknowledged the alternative way of learning offered by the test, 1 wrote “nothing”.

The question “What did you dislike?” was answered 53 times: 52 disliked the very restricted feedback of displaying only the number of correct answers (27 missed a full solution comparable to the geotechnical quiz in OLAT, 25 missed the information which question of the set had been answered correctly), 1 stated that the challenge of the test had been too high.

The question “What should be improved?” was answered 42 times: 21 requesting full solutions, 20 requesting at least the information which question of the set had been answered correctly, one requested a further development of the test, as the fundamental idea of the tool was very good.

2) *Interviews:* We invited all students to take part in a personal interview on the provided e-learning tools after the written examinations. We wanted to conduct about 10 interviews. The interviewed students should cover the whole range of ability classes, thus we wanted to select them by their performance in the written exam. As we worried about having enough volunteers to randomly select two interview partners from each of the five performance classes (1 to 5, compare Fig. 8), we established the following incentive scheme. All students who applied for a possible interview were remunerated with 1/4 grade point for the following written exam (from 16 achievable grade points). The randomly selected interview partners gained additional 3/4 points when they participated in the interview. We invited 122 students via OLAT. 44 students applied for a possible interview and 11 were selected. One of the selected students did not show up in the interview. 6 of 10 admitted that they would not have applied for the interview without the incentive scheme. In retrospective, the incentives proved to be necessary to attract enough students for a random distribution over the

five performance classes.

The 10 interviews generally confirmed the result of the preceding paper and pencil poll. However, they gave much more inside information on the way students used the new tool. Some of the students had been very frustrated by the restricted feedback (only the number of correct answers of the total number of questions was displayed after each test) and stopped working with the test. Others (less frustrated ones) were alarmed by their weak performance in the first test. They expected a much better result, which means that their initial self-assessment did not fit. These students had been encouraged to repeat the test over and over again to reach at least 80% of correct answers. Well-prepared students, simply checked their performance with one or two tests. One student could not cope with the test at all. The challenge of the questions was much too high for his learning level. He accidentally found out that some tests could be correct to a high degree (over 50%) without choosing any answer. The test software behaves like this due to the random choosing of the three displayed answers from a large pool so that the possibility of displaying only wrong answers was rather high. He then repeated the test about 30 times without choosing any answer. He stored screen shots of tests with a high number of correct answers and worked out standard solutions by comparison and guessing.

Some students were disconcerted by the button “send” which had to be pressed for evaluating the test. They worried about the anonymity of performing the online MCT.

We had been worried about the production of collected standard solutions for the MCT, as we expected a much higher learning-by-heart-scenario. The interviewed students claimed that they would not initially have used such sample solutions for answering the test questions. However, they admitted that they would have used them if they had not been able to answer the questions. Some of them would have tried to answer the failed test questions through repeated learning of the related topics, others (typically students with a lower degree of general

knowledge) would go straight for the sample solution. All students assumed that using sample solutions expedited the preparation for the written exam, but at the cost of drastically decreasing any fundamental understanding of the matter, as well as of a vanishing long-term learning effect.

3) *Interpretation and consequences:* Overall, both our hypotheses have been confirmed. However, it was more obvious for students to recognise the value of the online MC test for their exams than to realise the influence of the test on their calculation skills. The latter effect was realised by students more easily in the interviews.

Students would strongly recommend the test to other students, as they rate the positive effect of performing the test on their knowledge and they did not find the MCT to be superfluous, see Table. III. This may be due to the fact that they found the online multiple choice test more related to the exam than the geotechnical quiz and therefore more helpful, see Fig 10.

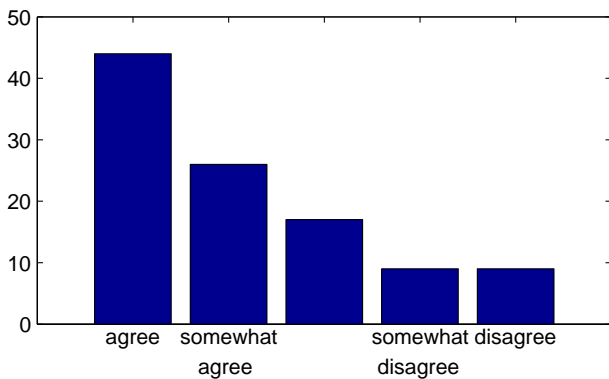


Figure 10. Rate of agreement to the statement: The MCT served as good preparation for the written exam.

Students generally tend to use such tools just before the examination which is confirmed by the event log of the software, see Fig. 9. The higher usage before the second exam may imply that students found out the relevance of the online MCT for the written exam and communicated this to other students, which was in line with the rather high degree of agreement to the statement "I would recommend the MCT to other students", Table III and Fig. 11.

Students were generally frustrated by the very restricted feedback of the multiple choice part, which may also be a reason for the only mediocre evaluation of the MCT. They often requested information about which question of the set was correct. We assume that this may be due to a misunderstanding of the purpose of the online MCT. This test was regarded as a self-assessing tool and not as an additional learning tool like the geotechnical quiz. Especially in the interviews, it turned out clearly that learning is not possible simply by repeating the online test over and over again. One has to go back to the "start": reading the lecture notes, related books, repeating the calculation examples and so on. By applying such repetitive steps, the follow-up test turned out to be more

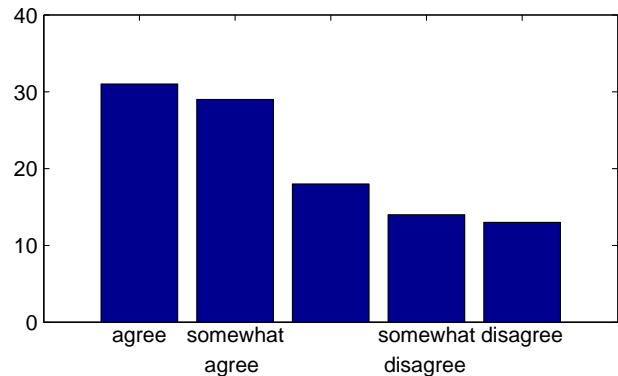


Figure 11. Rate of agreement to the statement: I would recommend the MCT to other students.

satisfactory. This interpretation is consistent with the fact that the number of performed tests did not relate to the performance in the written exams. To give a guide to topics that should be considered in the repeated learning loops, we changed the software to mark questions which were answered correctly. We realized that the production of sample solutions by the students would be much easier with that enhanced feedback. Our faith in the individual responsibility of the students for the proper use of such unofficial information was strengthened by the responses to that issue in the interviews. In the future, we will better communicate the purpose of the two e-learning tools, which was also suggested by some interviewed students.

The thoughtless test repetitions by one interview partner, to work around real learning, clearly showed a flaw in the software. Although we expected a strong decrease of such a behavior due to the enhanced feedback, we implemented a user-defined upper limit for questions for which only wrong answers are chosen from the answer pool.

The worries concerning anonymity when performing the test are taken seriously, as we want to enhance a free learning environment without any external pressure. We changed the send button to an evaluation button. In addition, we informed the students about the anonymity of both e-learning tools in the introduction, the handout and a newsletter.

C. Further applications

Recently, we have applied the online MCT as a pre-examination tool to oral examinations in three other lectures in our division. Such oral examinations are performed in groups of up to four students. Over the years, it has turned out that poorly prepared students lower the average grade of the whole group as the objectivity of a possibly annoyed examiner is limited, which in turn biases his overall impression of the group. A multiple choice pre-examination of factual knowledge should filter somehow better prepared students, thus reducing the examiner's assessment bias. The online MCT is evaluated automatically and saves a lot of time.

PAPER
MULTIPLE CHOICE TESTS: MORE THAN A TIME SAVER FOR TEACHERS

Table IV
EXAM GRADES

winter semester	number of tested students	mean grade	standard deviation	
2002	52	2,6	1,1	diploma curriculum
2003	61	3,1	1,1	
2004	61	2,9	1,0	
2005	52	2,9	1,1	
2006	64	2,7	1,2	
2007	79	3,9	1,0	commencement of bachelor program
2008	107	3,6	1,0	first group of bachelor students in course
2009	104	2,9	1,1	
2010	93	3,8	1,0	
2011	114	3,7	1,1	implementation of quiz and MC test part in the written exam
2012	114	3,8	0,9	
2013	106	3,0	1,1	implementation of online MC test

The Chair of Geotechnical Engineering of the Bauhaus University Weimar in Germany is currently implementing our OLAT geotechnical quiz in their open-source learning platform Moodle⁶.

IV. THE IMPACT OF THE ONLINE MC TEST ON FINAL EXAM GRADES

Exam grades cannot reflect any long-term learning effect. However, one can analyze a time series of such grades. Doing so, one has to bear in mind that the student population changes every year and the written exams most likely do not have the same level of difficulty. The evaluation of the time series of course grades in Table IV reveals a slight difference of students' performance in the old diploma program and the new bachelor program. Diploma students attend the course in their fifth semester, whereas bachelor students attend the course in their third semester, a year earlier. There was no change in the course content and the methods of teaching and of the examination. The calculation of the mean value of the time series from mean grades of 2002 to 2007 yielded 3.1 with a standard deviation of 0.5. The courses had been attended solely by diploma students. The bachelor program commenced in 2007 and the first group of bachelor students attended the course in their third semester (second year), i.e. 2008. The average of the mean grades from 2008 to 2013 is 3.4 (standard deviation 0.4), which is slightly worse than the diploma students (3.1). It is important to bring to attention that the change of the mean value is less than the standard deviation. However, it seems obvious that students attending a geotechnical course earlier in their studies are less trained in engineering basics (mathematics, mechanics) and are therefore more likely to perform poorly in soil mechanics.

The implementation of the geotechnical quiz in 2011 did not change the overall performance of the students, probably due to the simultaneous change of the written exam, which included a MC test part. The students were unfamiliar with the MC test and performed rather poorly. Therefore, any positive effect of the quiz may be concealed by the change in the exam.

⁶<https://moodle.org/>

The implementation of the online MC test in 2013 raised the performance of the bachelor students up to a level similar to the mean performance of the diploma students. The difference of the mean of 2008 compared to 2013 (3.4) is just equal to the standard deviation of the time series (0.4). However, the trend indicates a positive effect of the tools which seems to alleviate the effect of the course relocation from the fifth semester to the third semester.

V. CONCLUSION

The development of the tests required a tremendous amount of work. Feedback in such online tests is very important, which can gradually be reduced from a learning tool like the quiz to a self-assessment tool like the online multiple choice test. We are convinced that the integration of such online tests was successful with respect to student activity and long-term training effect.

ACKNOWLEDGMENT

We thank Franz Kroath for his didactic guidance, the students Hannah Göttgens, Michael Huber, Matthias Öhlböck, Andreas Vigl und Hannah Wolf and our colleagues Chien-Hsun Chen and Barbara Schneider-Muntau for their intensive cooperation. Further we thank Franz Kroath and Iman Bathaeian for their linguistic assistance.

REFERENCES

- [1] "Online learning and training." [Online]. Available: <http://www.olat.org/>
- [2] E. J. Marsh, H. L. Roediger, R. A. Bjork, and E. L. Bjork, "The memorial consequences of multiple-choice testing," *Psychonomic Bulletin & Review*, vol. 14, pp. 194–199, 2007.
- [3] M. A. Lundeberg and P. W. Fox, "Do laboratory findings on test expectancy generalize to classroom outcomes?" *Review of Educational Research*, vol. 61, no. 1, pp. 94–106, 1991. <http://dx.doi.org/10.3102/00346543061001094>
- [4] E. J. Palmer and P. G. Devitt, "Assessment of higher order cognitive skills in undergraduate education: Modified essay or multiple choice questions?" *BMC Medical Education*, vol. 7, p. 49, 2007. <http://dx.doi.org/10.1186/1472-6920-7-49>
- [5] H. L. Roediger and E. J. Marsh, "The positive and negative consequences of multiple-choice testing," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 31, no. 5, pp. 1155–1159, 2005. <http://dx.doi.org/10.1037/0278-7393.31.5.1155>
- [6] A. C. Butler and H. L. Roediger, "Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing," *Memory & Cognition*, vol. 36, pp. 604–616, 2008.

PAPER
MULTIPLE CHOICE TESTS: MORE THAN A TIME SAVER FOR TEACHERS

- [7] E. J. Marsh, J. P. Lozito, S. Umanath, E. L. Bjork, and R. A. Bjork, "Using feedback to correct errors made on a multiple-choice test," *Memory*, vol. 20, pp. 645–653, 2012.
- [8] H. L. Roediger, P. K. Agarwal, S. H. K. Kang, and E. J. Marsh, "Benefits of testing memory: Best practices and boundary conditions," in *New frontiers in applied memory*, G. M. Davies and D. B. Wright, Eds. Brighton, UK: Psychology Press, 2010, pp. 13–49.
- [9] H. L. Roediger, A. L. Putnam, and M. A. Smith, "Ten benefits of testing and their applications to educational practice," in *Psychology of learning and motivation: Cognition in education*, J. Mestre and B. Ross, Eds. Oxford: Elsevier, 2011, pp. 1–36.
- [10] M. Schneider and M. Mustafić, Eds., *Gute Hochschullehre: Eine evidenzbasierte Orientierungshilfe*. Heidelberg: Springer, 2015.
- [11] Division of Geotechnical and Tunnel Engineering, University of Innsbruck, "Collection of examples of written examinations." [Online]. Available: <http://www.uibk.ac.at/geotechnik/stud/pruefungs-fragen.html>
- [12] W. Fellin, *Geotechnik – Lernen mit Beispielen*, ser. Advances in Geotechnical Engineering and Tunnelling. Berlin: Logos, 2003, vol. 7.
- [13] D. J. Elton, *Soils Magic*, ser. Geotechnical Special Publication. American Society of Civil Engineers, 2001, no. 114.
- [14] I. Herle and S. Gesellmann, "Demonstration experiments in geotechnical education," in *Education and Training in Geo-Engineering Sciences*, I. Manoliu and N. Radulescu, Eds. CRC Press, 2008, pp. 379–382.
- [15] J. C. Marques, "Experimental, numerical and virtual tools for civil engineering students," *International Journal of Engineering Pedagogy*, vol. 1, no. 1, pp. 42–45, 2011. <http://dx.doi.org/10.3991/ijep.v1i1.1605>
- [16] S. Alam and L. Jackson, "A case study: Are traditional face-to-face lectures still relevant when teaching engineering courses?" *International Journal of Engineering Pedagogy*, vol. 3, no. 4, pp. 9–15, 2013. <http://dx.doi.org/10.3991/ijep.v3iS4.3161>
- [17] G. Medicus, "Fördert ein Multiple-Choice-Selbsttest das Verständnis für geotechnische Problemstellungen?" *Schaufenster Lehre*, Universität Innsbruck, 2014. [Online]. Available: <http://www.uibk.ac.at/rektorenteam/lehre/die-lehre-seite/schaufenster/publikationen/2014-10-02-medicus-lehrgangprojekt-20141002.pdf>
- [18] B. Jacobs, "Aufgaben stellen und Feedback geben," 2002. [Online]. Available: <http://psydok.sulb.uni-saarland.de/volltexte/2004/438>
- [19] —, "Testfragen selbst beantworten oder Musterlösungen studieren?" 2010. [Online]. Available: http://psydok.sulb.uni-saarland.de/volltexte/2010/2693/pdf/testen_vs_musterloesung.pdf
- [20] R. L. Bangert-Drowns, C.-L. C. Kulik, J. A. Kulik, and M. Morgan, "The instructional effect of feedback in test-like events," *Review of Educational Research*, vol. 61, no. 2, pp. 213–238, 1991. [Online]. Available: <http://rer.sagepub.com/content/61/2/213.abstract>
- [21] D. Kolymbas, *Geotechnik : Bodenmechanik, Grundbau und Tunnelbau*, 3rd ed. Dordrecht, Heidelberg, London, New York, Berlin: Springer, 2011.
- [22] T. Hell, W. Fellin, and G. Staggl, "Webbasierende Prüfungsaufgabendatenbank zur studentischen Vorbereitung auf Multiple-Choice-Tests," eLearning Projekt, Universität Innsbruck, 2012.
- [23] H. Altrichter and P. Posch, *Lehrerinnen und Lehrer erforschen ihren Unterricht. Unterrichtsentwicklung und Unterrichtsvaluation durch Aktionsforschung*, 4th ed. Bad Heilbrunn: Klinkhardt, 2007.

AUTHORS

W. Fellin is with the Division of Geotechnical and Tunnel Engineering of the University of Innsbruck, A 6020 Innsbruck, Austria (email: wolfgang.fellin@uibk.ac.at)

G. Medicus is with the Division of Geotechnical and Tunnel Engineering of the University of Innsbruck, A 6020 Innsbruck, Austria (email: gertraud.medicus@uibk.ac.at)

The work was supported in part by the e-learning projects 2012.249 and 2012.273 (University of Innsbruck), as well as by the office of the vice-rector for student affairs and teaching, University of Innsbruck.