The ten pitfalls of digital testing

Metha Kamminga
June 2011

Experiences at NHL University of Applied Sciences with the digital testing system Maple T.A. (Testing and Assessment) have been described before in *Euclides*. Digital testing continues to develop; it is therefore time to take another look at the current state of affairs.

The sophisticated testing system Maple T.A. is especially suited for testing knowledge and skills in mathematics- and technical courses thanks to the various possibilities of randomization (including textual randomization) and the use of formulas and plots. The system can be perfectly well applied in other areas as well, without any loss of possibilities. Use of this system has become widespread: it has undergone a great many improvements and is now offering a very rich environment for digital testing.
Recently, I have been paying a great deal of attention and offering support to various teams of teachers who have started to apply digital testing with Maple T.A. to assist their courses, not just at NHL University, but throughout the country. This has enabled me to point out a number of pitfalls due to which the expected success often remains forthcoming. Understandably, this leads to frustration both with teachers and students. Digital testing is more and more applied at Universities of Applied Sciences and Research Universities to stimulate studying and guide the learning process, not just to lighten the teacher's workload. However, it does demand a certain outlay of time and effort that amply repays itself in time saving, better results of, and a more congenial contact with, students.

This article deals mainly with formative testing used to assist the educating process. What is involved here is not the decision to have a student pass or fail the test, but to guide him through the curriculum and have him adopt a critical attitude with which to manage his own learning process. It is also perfectly possible to support the route to the so-called competency test with knowledge- and skills tests in which feedback and supervision play a major part. It need not necessarily concern only the subject Mathematics in Higher Vocational Training. The experiences may quite easily be translated to any other subject in any other sector of education.

Pitfall 1 - The entry test

Having incoming students sit a digital summative test without preliminary training in the digital testing system is the very first pitfall. Often we, the teachers, know the outcome of such an entry test in advance; the students will be instantly discouraged by an insufficient score, even though it gives a lot of information about the gaps in their knowledge and they won't be graded. After this, getting them fired up again for the subject and have them take up digital training will prove tricky. So no entry test, not even a written one, but rather an *entry track* concluded by a end test either digital or not. This end test can be taken digitally as a diagnostic test to identify the subjects not (yet) mastered by the student after having completed the entry track.   The Maple T.A. testing system offers a lot of possibilities as regards the analyzing of such an end test. Subsequently, depending on the outcome, a follow-up track may be offered.

Pitfall 2 – Extensive practising tests

Frequently we see that the digital practising tests prepared for the students are too long and too extensive. Students whose score is insufficient will feel reluctant to do the test again when it takes up too much time. Much better to offer short tests in the form of a 'ladder', by which we mean a gradual building up by offering a chain of short tests (15 minutes) in a certain order and with gradually increasing difficulty. The system makes it possible to bar the student from the next test unless he has completed the previous one with sufficient results. Here, the randomization- and feedback possibilities of the testing system come into their own in a big way, especially in the field of mathematics. You might want to number the tests; this makes it easier for students to see how far they have progressed on the ladder. Don't make the tests too long; 15 minutes should be sufficient.

Pitfall 3 – Offering an unlimited number of attempts.

As was pointed out earlier in an article dated June 2008 many years of experience have proved that with offering practising tests the ideal maximum number of attempts at any one test is four. After four attempts with insufficient results the student must absolutely be stopped to prevent any misconceptions from being formed. This is the time for teacher or assistant to find out what the problem is. The testing system offers possibilities to give additional feedback. Of course, it is even better, together with the student to look in his exercise book to try and detect where it keeps going wrong; if necessary, an extra attempt may be offered. A good number of teachers are already adapting this maximum number of attempts to be set and are reporting the same experience with the magical number of four: the 'frustration turning point.' A student knowing beforehand he will be given a maximum of four attempts at a test will handle the system better; after one or two failed attempts he will be ready to delve into the course material first before the next attempt.

Pitfall 4 - A six is a pass

The criteria for passing a practising test should not be too low. A 60% passing score is far too low for a teacher to assume that the course material has been understood and the skills have been mastered. Mistakes due to sloppiness should be weeded out; thus, the student learns to work accurately. Skipping 'difficult' questions will occur less frequently. Even subject material that uses digital multiple choice tests still allows the scoring of points by guesswork; a 60% score is absolutely no guarantee for having mastered the course material. Surprisingly, students usually accept the heightened minimum score of, for instance, 80% readily enough. An added advantage is that in the end test where a six actually means a pass, the scores are generally higher than in former times when the learning process was not yet accompanied by digital testing. In addition, it gives students' morale a boost when, in their grades lists, they see high grades such as eights, nines and tens. Also, if a student has, for instance, scored 80%, he will definitely want to know what that one mistake was. This means that you should always translate the score into a grade; give ten points for a practicing test so that with an 80% score an eight will show up in the grades list. A rock-solid morale booster!

Pitfall 5 – Omitting to analyze the test results

One should not expect digital testing to take all the work out of one's hands completely. It is important always to analyze the test results at item level. Determining the degree of

complexity might have to undergo some fine tuning after one has analyzed the test results at item level. Which items result in bad scores and why? Which items generate a lot of questions? Are these items not too difficult relatively speaking? Do they fit into the streamline of the consecutively offered exercises? Sometimes the language used is too unfamiliar for students which necessitates the rewriting of items or complementing them with useful hints and, where needed, feedback. We have noticed that many concepts students are not familiar with, both in mathematics and in other subjects, may become an obstacle for their performance in the tests. You might consider occasionally throwing in a little 'language test' among the consecutive practising tests, in which students have to fill in the correct words in the sentences offered. This will work for any subject including mathematics. It forces students to stop and ponder the concepts they are unfamiliar with, thus helping them to use them correctly in a given context. In mathematics, think of concepts such as factors, factorizing, coefficients, powers, terms, exponents, functions and equations.

Pitfall 6 - Non-obligatory participation

When we first started offering digital tests we fell headlong into pitfall number 6, that of non-obligatory participation. Good students cheerfully started practising; students who really needed it, did not. As a consequence, the gap between good students and the less gifted ones became increasingly wider. Sitting tests is always fairly threatening and after scoring the first insufficient the fun soon goes out of it.

Digital testing requires a firm hand. Making the practice track obligatory has been readily accepted; since it produces results, not a single student is against it. There is however one condition: the track must incorporate the possibility of differentiation as regards level and steepness of the learning curve. In the past we definitely underestimated the need for this, which resulted in a stream of complaints when we made running through all the exercises and scoring at least 80%, obligatory. The good students didn't see the need for all that exercising, and rightly so, whereas the not-so-good ones frequently still got too little practice and thought everything too difficult. At a certain moment we started offering two tracks: a fast one for students who were able to proceed through the exercises with seven-league strides, and a slower one in which the others had to do all the exercises. Last year we even introduced a third track. A slow track in which all practising tests are obligatory, the starting level is slightly lower and the learning curve flatter; also room has been made for marking time, in order to be able to practice. Next, the regular track with a slightly steeper learning curve and finally, the fast track in which only the final tests of each subject are obligatory. You will be interested to know that students are allowed to assign themselves to a track and are quite happy to do so. In practice, over 50% of the students opts for the slow, and in their eyes, safe track. This is a somewhat misleading idea, since everybody has to have completed all the exercises at the same time before the date of the final test. In the first three weeks of the course students are allowed to switch tracks themselves, but after three weeks they know perfectly well which track fits them best. Movements up ánd down do occur. Switches later than that require the involvement of teacher or assistant. Once a track has been chosen all tests belonging to it must be passed before one can obtain an 'entry ticket' for the final test. We started the current academic year by adding an extra-slow fourth track to the ones already existing. However, we have now reduced the number of four back again to three, which in our opinion is best for the present circumstances.

Digital testing stands or falls with obligatory participation ánd a firm hand ánd the possibility of differentiation.

Pitfall 7 – Bonus points

Many teachers are accustomed to give bonus points to students who have gone the extra mile in their subject. In the digitally controlled practising track it is important not to give out any bonus points at all. Years ago we fell into this trap in a big way. We saw that giving out bonus points increased the risk of fraudulent behavior. So no more bonus points; instead, an obligation to complete all the exercises belonging to a learning track! In practice it turns out that students do their own tests anyway because they want to be confident of success and show up well in the end test. The randomization option in the testing system, too, helps to prevent students from copying each others' answers. It is really surprising to see how eager they are to participate in the final testing even though no points were granted beforehand. And so the number of participants passing the end test stands, at present, at 90%.

Pitfall 8 – Setting time limits for tests

The system allows a time limit to be set for tests. For digital end tests this is ideal. How long a student is allowed to work at the end test must be accurately calculated beforehand. During the test the time remaining is shown on screen; when the time has run out the only action still possible is clicking on *grading,* after which the score is displayed. Also for diagnostic tests, within a protected environment and with supervision, it is a good idea to set a time limit. However, with practicing tests that students do in their own time it is absolutely imperative, for various reasons , *not to set a time limit!* Above all, this has to do with students working at the practising tests in their own time. This means they may be sitting somewhere very crowded or where they may be disturbed by telephones or whatever. Skipping through the test or sloppy mistakes due to haste will all lead to frustration, since a minimum score of 80% is a requirement. The most important thing is to prevent frustration from occurring and to offer a safe environment.

The beauty of the Maple T.A.-system is that with no time-limit having been set one can always stop halfway down the test. All the filled-in answers are saved and one can complete the test at a later time. Also, when the student becomes aware at first glance that he is not yet ready for the test, he may want to study the learning material first. The test may be completed at a later time. The student works, peacefully and quietly, at raising his level, has time to ask questions and study the learning material once again. So the motto is: no time limit on practicing tests!

Pitfall 9 – Communication

Digital testing in support of a course may have a time-saving effect for the teacher. But it would be a mistake to assume that by using digital testing everything will be taken care of by the system. True, the piles and piles of correction work are a thing of the past, but something else takes their place. Through communication with the student a safe environment has to be created in order to prevent frustration; this can be done in various ways. If it is at all possible, have students avoid the use of emails; if you don't there will be no end to them. That, too, is a pitfall. Usually, there is a digital learning environment available which enables students to ask questions in a 'digital classroom'. When a course is 'running', it is necessary to check for questions regularly and by doing so ensuring good progress. During class, difficult problems can be dealt with, exercise books checked and progress monitored. So you have to make sure that the line of communication between students and teachers (and any assistants) is open, clear and takes place at regular intervals.

Pitfall 10 – Omitting to supervise students' handling of the testing system

It would be wise to put some extra effort in supervising the handling of the testing system at the start of the course. By doing so, students will be getting the hang of it more speedily. Have students cast a critical eye at the outcomes of the tests. During the exercises, give as much feedback as you can, coupled with pointers and in-between steps. Not only students have to learn how to handle the system; teachers and assistants, too, should not fall into the trap of thinking that everything will take care of itself.
It is possible to see in the system which students use up many attempts for each exercise. One should advise these students to assign themselves to a slower track; alternatively one might want to check whether they are using the correct procedures.
When students have used up their four attempts and still have failed to score, it is the first task of teachers and assistants to try and find the cause of these failures, before they allow yet another attempt. Also teachers and assistants should train themselves to add feedback to the reassessment; students must not be misled into thinking that something that is not entirely correct is assessed as correct when no further comment is added.
And lastly, if a student should fail the end test, it is important to try and find out what caused it; after all, the student has completed the whole practising track. Always, in such cases, it is possible to find the cause and very often it turns out that is has to do with handling the testing system. Possible diagnoses are: the student has omitted to have a good look at the feedback; he has omitted to cast a critical eye on mistakes made; he has failed to read the questions properly; he has drafted the exercises and failed to write them down in a notebook where mistakes can be properly marked and corrected; the student finds himself in the wrong track, has had too little practice as a result, which has escaped the teacher's notice; the student has had too much 'help' while making the exercises. And so I could go on, listing possible causes for failing the end test in spite of having successfully completed the full practising track.

Some pointers for mathematics and related subjects

- In mathematics and related subjects preference should be given to written summative tests over digital ones, because with the former it is possible also to examine the way students have arrived at their answers.
- All assignments must be completely worked out in a notebook not as drafts.
- After the test the student will examine and assess the notebook himself. Teachers and assistants will check the notebooks.
- Limiting the use of calculators can be easily enforced in the system by using exact answers or by inputting the calculations themselves.
- Attention should be paid to the lay-out of the questions appearing on the screen .
- Equally important is indicating clearly how each answer should be filled in.