Digital Testing in Engineering Courses

The use of scenario questions in the digital test system, Maple T.A.

- **Celine Goedee**
  Test Expert
- **Meta Keijzer-de Ruijter**
  Expert digital testing
- **Erik Offerman**
  Teacher and Researcher

All three work at TU Delft. This pilot was sponsored by 4TU – Centre for Engineering Education.
E-mail: M.A.Keijzer-deRuijter@tudelft.nl

At TU Delft, a lot of courses have exams that require the execution of complex calculations. The most heard argument for not having digital exams is the restriction that students can only enter their final answer and that they cannot show the steps they took to get there, thus missing out on a possible partial score for the problem. By creating 'scenario' questions, this restriction can be addressed.

Introduction

In the test system Maple T.A., it is possible to use so-called scenario questions. At TU Delft, we found that the problem with using this type of question was that we had too little experience to give out guidelines on how to create qualitatively good scenario questions. This is why we conducted a pilot to focus on creating guidelines, instruct students and measure the quality of the assignments. In this paper we first describe the two main types of scenarios, followed by a step-by-step development of the questions. Finally the pros and cons of the scenarios are given and we make recommendations for working with these scenarios in the daily practice of testing.

Description of the scenarios

Simply said, a scenario question is a digital test question consisting of multiple, inter-related questions that are presented consecutively. First, the main question is presented and after answering, the student is presented with sub questions. The teacher defines the point distribution between the main and sub questions and the number of attempts for the different steps. We differentiate two main scenario types: Underpinning and Scaffolding.

Underpinning scenario

Figure 1 represents the underpinning scenario. In this scenario it is not just the correct answer that is of importance, but also the underlying reasoning or calculations. In the first section of the problem the student answers the main question.

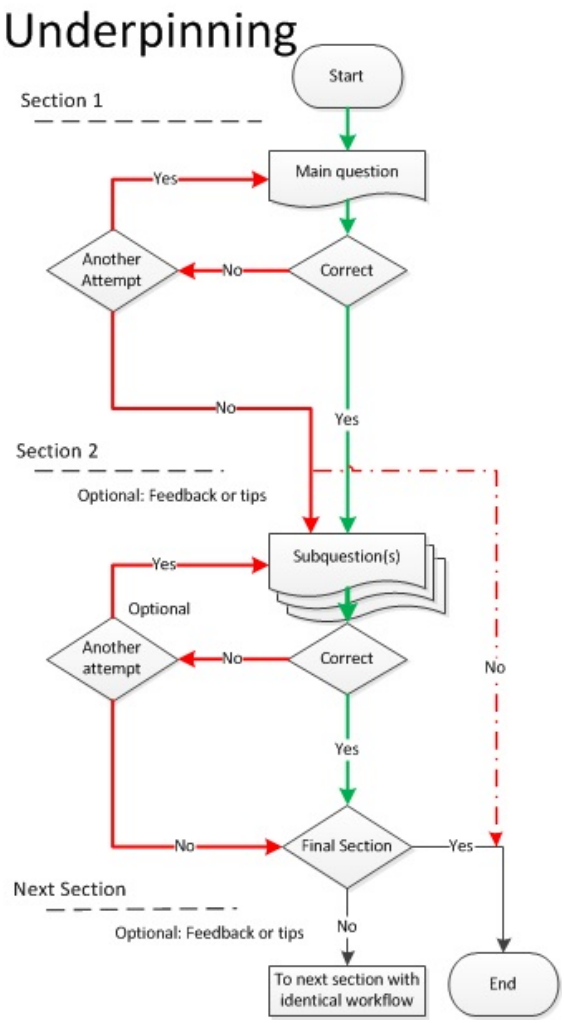


Figure 1: Flowchart of the underpinning scenario

By clicking the “verify”button (see figure 2) the computer checks the student’s response. When incorrect, the student gets another opportunity to answer the main question (provided that another attempt is permitted).

When the main question is answered correctly, the student is guided to the next section including one or more questions that should underpin the student’s response to the main question. When, after multiple attempts, the answer to the main question is still incorrect, the system presents the sub questions or the main question is closed off (see the example in figure 2). The system can handle multiple sections that are built in the same structure. Also, the teacher can choose the number of attempts for each section.

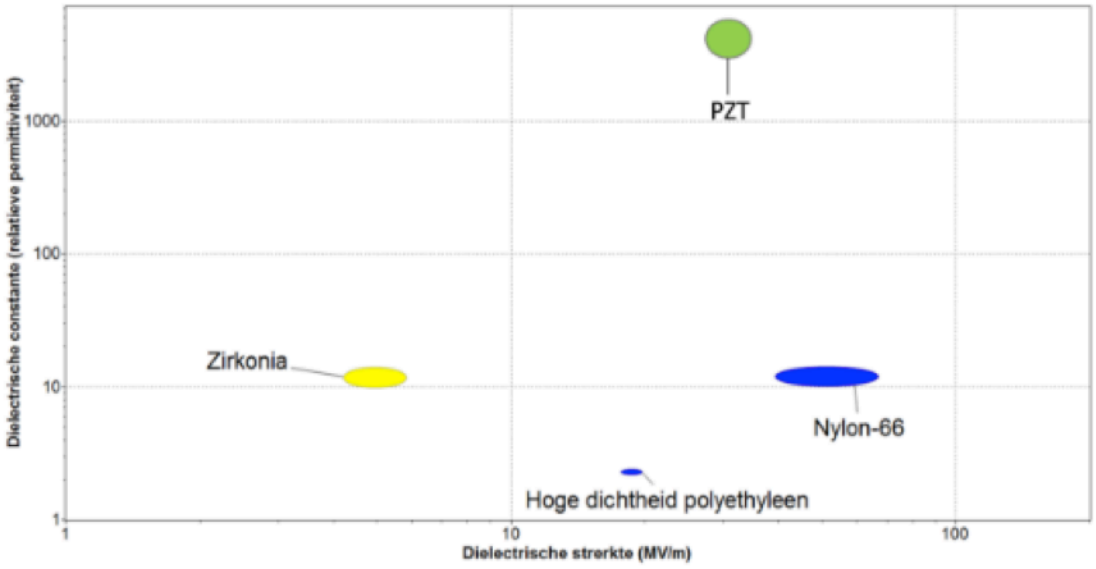
A scenario question is a digital test question that consists of multiple, inter-related questions

In figure 2, an example is presented that contains a multiple choice question in the main section. In this case the teacher chooses to close the problem when the main question is answered incorrectly. As stated at the top of the question page, students need to correctly answer all questions in order to gain the full score for this problem.

In figure 3, an example of an open, automatically graded, main question is given. In this case, the problem requires a well-defined procedure to solve. In order not to present this procedure right at the start, an underpinning scenario is chosen. Regardless of whether a student correctly or incorrectly answers the main question, all students need to go through all sub questions.

After correctly answering the main question you will get 3 additional questions, to corroborate your choice. All questions must be correct to get a score on this question.

A supercapacitor can be used to store electrical energy. This requires a material with dielectrical properties. In the graph below the dielectrical constant ε_r is presented versus the dielectric strength E_b (MV/m) for four different materials.



Which dielectrical material has the highest energy density E (kJ/m^3) for the storage of dielektrical?

(Click for List) ▾

Section Attempt 1 of 1

Verify

After clicking the verify button the sub questions appear. Answering these should corroborate the choice

a. Give the materials index that is maximised to find the material with the highest energy density.

$M =$

Click here for an overview of the Greek Symbols and how to write them

b. What is the slope of the line of constant material index ?

Number

c. In what direction should this line be shifted to determine the maximal material index?

(Click for List) ▾

Section Attempt 1 of 1

Verify

Figure 2: Example of a multiple choice underpinning scenario

After answering the main question you get 3 additional questions to corroborate your response.

A long, thin-walled, cylindrical tube with a given length L (m) and inner en outer radius R_i (m) and R_u (m), respectively, is exposed to a force F (N). Both ends of the tube are supported as shown in the figure below.

We aim to minimize the material cost under the constraint that the tube's outer fiber is not allowed to deform plastically due to the bend momentum M_{buig} (Nm).

The tube's wall thickness t (m) is, to a certain extend, the free parameter of this system. The only restriction is that the wall thickness is far less than the inner radius of the tube: $t << R_i$. The following simplification can be used:

$$R_i + t \approx R_i \text{ and}$$

$$R_u^2 - R_i^2 \approx 2 \cdot R_i \cdot t$$

The material the tube is made of has the following properties: density ρ (kg/m^3), yield stress σ_y (Pa) and cost per kilograms C_m (€/kg).

Main question

Derive the material index that has to be maximized to select the right material for minimal material cost C for the forementioned situation and constraint.

Material index =

Click here for an overview of Greek symbols and their spelling

Section Attempt 1 of 2

Verify

After clicking the verify button the sub questions appear. Answering these should corroborate the choice

a. Give the formula for the target function:

C =

b. Give the formula for the constraint:

M_{buig} =

c. Derive the equation for the free parameter:

Give yhe symbol for the free parameter in the first cel and the formula for the free parameter in the second cel

=

Section Attempt 2 of 2

Verify

Figure 3: Example of a numerical question underpinning scenario

Each of the questions represents a partial score, adding up to a full score and allowing for a partial score. Sub questions are presented in separate sections, so a mistake in one question does not necessarily mean that consecutive questions should also be graded as being incorrect. In figure 4, the scaffolding scenario is presented schematically. This scenario emphasizes the ability to correctly solve the main question, regardless of the way to solve it.

In the first section of the problem, the student answers the main question and verifies the response. When the main question is answered correctly, the student receives the full score and no sub questions are given. When, even after multiple attempts, the response to the main question remains incorrect, the student receives sub questions to enable him or her to show the level of competence in smaller steps. In this way a partial score for the problem can be achieved. Figure 5 shows that in the second section an intermediate question is posed, after which the main question is repeated. Two attempts are allowed for both the main question and the sub question.

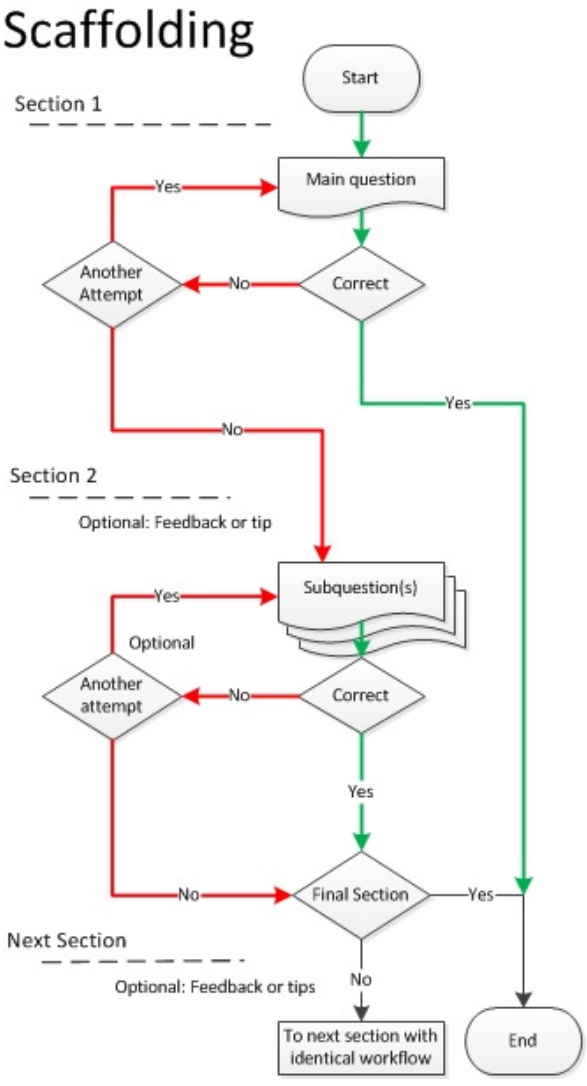


Figure 4: Flowchart of the scaffolding scenario

Pilot

In the pilot, we took a closer look at the quality of the scenario questions. Firstly, it was expected that students would wield an assessment strategy if they were given multiple attempts for answering the questions and a certain score distribution between main and sub questions. Secondly, we investigated whether the two different scenarios have an added value.

During the pilot, we collected data in multiple ways: instructor input, student interviews, think aloud protocol with ten students taking a small test consisting of three questions, results and analysis of the exam (n=451) and remarks from the course evaluation. In this article, results and recommendations are given related to the steps in the assessment cycle.

Step 1: Construction of main and sub questions

Two boundary conditions are applicable to a scenario question: Solving the problem requires multiple steps and the response must be graded automatically. To gain insight into the problem, the following questions helped the instructor to pose the question in the 'digital form':

- What level of knowledge do you want to assess?
- What focus do you want: on the correct use of the procedure or getting the correct response to the main question?

When you answer the main question incorrectly, you get 2 sub questions

The microstructure of a hot rolling carbon steel contains microchemical bands of silicon (Si), as shown in the picture to the right. The light and dark parts in the picture are areas with resp. high and low concentration of silicon that is substitutionally resolved in iron.

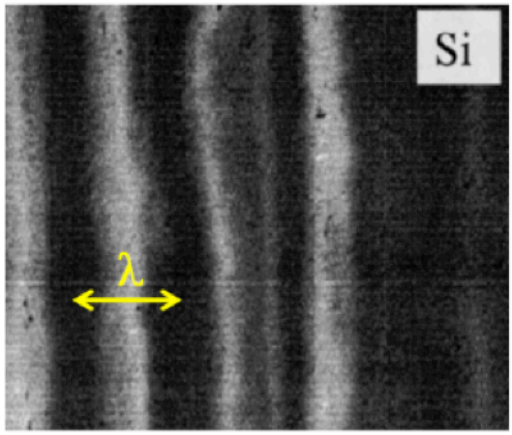
The mean distance between the bands, λ, is 100 μm .

The pre-exponential factor for diffusion of Si in iron, D_0, is $2.1*10^{-4}$ $\frac{m^2}{s}$

The activation energy for diffusion of Si in iron, Q_d, is 252 $\frac{\text{kJ}}{\text{mol}}$.

To obtain a more homogenous distribution of the Si-atoms this steel was heated to a temperature of $T = 1200°C$.

Assume the Si-atoms need to diffuse over a mean distance of ½ λ= 50 μm in order to obtain a more homogenous distribution of SI atoms over the material.



Main Question

Calculate how many hours the steel material must remain at a temperature T of 1200°C to obtain a more homogenous distribution of SI atoms over the material.

hours

Write your answer in 2 significant figures.

Only students that have answered the main question incorrectly - after 2 attempts -will get the sub questions.

Section Attempt 1 of 2

Verify

a. Calculate the diffusion-coefficient of Si in iron at $T = 1200°C$.

$\frac{m^2}{s}$

Write your answer in 2 significant figures.

b. Calculate how many hours the steel material must remain at a temperature T of 1200°C to obtain a more homogenous distribution of SI atoms over the material.

hours

Write your answer in 2 significant figures.

Section Attempt 1 of 2

Verify

Figure 5: Example of a scaffolding scenario

- Are there different ways to solve the problem? If so, do you want to see one specific way?
- What are the common mistakes students make?
- Which scenario (scaffolding or underpinning) suits your situation best?

For inspiration for the main question, we looked at the written exam version of the problem, the learning objective and the final question of the step-by-step questions in the written exam version. The sub questions are inspired by the detailed elaboration of the problem.

Step 2: Student preparation

Students in our pilot already had some experience in using the TU Delft test system, Maple T.A. The basics of navigation were not

an issue but the use of features yet unknown to them took them an unnecessary amount of time and sometimes resulted in unintentional incorrect entries. For instance, when entering formulas using the equation editor, they had trouble entering symbols in the correct location (subscripts, superscripts, etc.). They preferred using the 'text entry mode' in combination with the formula preview. After this adjustment, spelling of the Greek alphabet turned out to be a problem (for example writing 'labda' instead of 'lambda'). Thus it is important to identify the specific potential problems students might run into and give them the possibility to properly prepare for the exam. In the pilot, students got a practise test where ambiguities of the system were presented. In addition, in the lectures there was ample attention to the correct syntax for putting entries into the system.

It is still necessary to review automatically graded responses

Step 3: Taking the test

The following recommendations on test taking are assembled from the 'think aloud' sessions with students:

1. When looking at an image in the test, one student made a remark that the image was probably irrelevant, since it was rather small and hard to read. So make sure that the possibility to enlarge an image is provided in the test.
2. When entering a numerical response students were in doubt about the number of decimals to enter. Would the computer grade their response as incorrect if they put in too many or too few decimals? Thus, specify the accuracy that is expected from them.
3. Text written right underneath a response field is very often overlooked. Place relevant remarks about the entry expected above the response field.
4. Furthermore, two points occurred that the instructor needs to be aware of and tackle where possible:
    a. Students want to be able to browse through the exam at the start. When using the scenario questions, they can only see the main questions.
    b. Due to the fact that the student must verify his or her response, the mind-set will be influenced: They can either feel stressed or become motivated.

Step 4: Reviewing the exam

It is the experience of the instructor that a review of the automatically graded questions is still necessary. Students make small mistakes that were not anticipated in the design of the question. This causes the response to be graded as incorrect, while in

some cases a partial credit would be appropriate. Besides, it is very valuable to gain insight into the mistakes students make. These insights allow for improvement of the next exam, by being able to better anticipate student mistakes and possibly program partial grading for these mistakes ahead of time. It is essential to document the reason for regrading a response. Students have the right to review their exam and ask the instructor to explain the grade.

Step 5: Analysis

The exam eventually contained four scenario questions and two regular numerical questions. The psychometrical analysis showed that these questions differed in P-value (as expected). The Rir- values (Point biserial) indicated that the scenario questions discriminated well between the good and bad performing students.

Test taking strategies

In this pilot, test-taking strategy is defined as a tactic used by the student to decide how seriously he or she will make an attempt on answering the main question based on other factors than actually mastering the topic.

Number of attempts

The test system can offer multiple attempts on answering the question. During the think aloud protocol, the students would use the attempts lightly as many times as possible, saying: “I still have plenty of attempts”. This resulted in the loss of precious exam time. One student noted: “The fewer attempts you are given, the more valuable they are. It forces you to give it your best effort”.

In the exam, we analysed how students used their second attempt. Mostly, it was used to make small changes (for example: rounding off) or checking their calculations. Correction on syntax was not applied very often (remark: a preview of the response is available prior to verifying, checking for syntax mistakes). Not all students used their second attempt.

The conclusion is that students indeed use a test-taking strategy based on the number of attempts given. Our recommendation: as an instructor, please carefully consider for each question what the students will use their additional attempts for. Do not allow for more than two attempts in a time constrained test, otherwise students will spend too much time on that question.

Distribution of points between main and sub questions

We expected that students would skip the main question, if they could see that they could only score a small fraction of the total score. In the think aloud protocol, it turned out that students hardly took notice of the point distribution. Their reaction was that the score was not of interest, because they wanted to make a serious effort on the main question regardless. Furthermore, they stated that it was “a waste of time” to read the text on the point distribution at the top of the page.

Concluding: We did not find an appropriate test-taking strategy based on point distribution. Students might develop such a strategy in the future, once they are used to these type of exams.

Reactions of students on scenario questions

The sub questions were set up as steps towards the correct solution of the problem. Some students acknowledged this structure by stating that these sub questions are small steps towards the correct answer. But not all students said so. They stated that still only a “final” (numerical) response was graded. Some students claimed that the course topic is not suitable for digital testing.

Pros and cons

Advantages

1. By first posing the main question, the proficiency of the student to answer the question on their own can be evaluated.
2. In case of a multiple choice question, a correct answer through an 'educated guess' or just blind guessing no longer suffices, since the grading can be programmed in such a way that all main and sub questions should be answered correctly to gain the full score.
3. In the scaffolding scenario, the students are free to use their own strategy to solve the problem.
4. Students can show their proficiency in smaller steps.

Disadvantages

1. The instructor must consider very carefully what response can be given by students in the main question section. Making a mistake here can have weighty and undesirable consequences.
2. In the underpinning scenario, if a student's correct answer is graded as incorrect, due to programming issues, he or she will be presented with the sub questions. Thi smight cause confusion when he or she uses a completely different strategy to answer the main question.
3. Students find it unfavourable that there is no option to correct their response to the main question once the sub questions are shown.

Conclusion

Scenario questions can provide a solution for courses in which it is important to test whether students can come up with the correct strategy and response without providing them with the steps at first. From the pilot, we learn that scenario questions can function very well. It is very important that students receive proper training to work with the system's navigation, syntax and peculiarities. Furthermore, there should be careful consideration on the question set-up and correct programming of the grading process, to avoid problems during the exam or when reviewing the results. Unfortunately, scenario questions cannot be used for (sub) questions that require manual grading.

This article is a translation of a publication in the Dutch periodical 'EXAMENS' titled:

Digitaal toetsen van engineering Lesstof.

May 2017