Developing Concept Inventory Tests for Electrical Engineering (CITE): extractable information, early results, and learned lessons

Eva Fjällström *, **Steffi Knorn** **, **Kjell Staffas** ** and **Damiano Varagnolo** *

* *Luleå University of Technology, Luleå, Sweden*
** *Uppsala University, Uppsala, Sweden*

Abstract: This paper suggests a method for developing, implementing and assessing a concept inventory test for electrical engineering students (CITE). The aim of this test is to help students better understand and learn core concepts, plus increase their awareness about links between the different courses and other themes of the program. Our and other experiences show that students often struggle to understand and use fundamental concepts, and how these relate to the various courses. This issue is probably due to the fact that traditional exams mainly focus on assessing procedural tasks (e.g., directly solving specific problems following step-by-step approaches). The investigated programs at Uppsala University (UU) and Luleå University of Technology (LTU), nonetheless, have no tool for collecting quantitative data on how students develop conceptual knowledge throughout the programs, and thus no means to obtain an holistic view about their learning process. The here proposed methodology thus describes how to develop tests that would not only provide students with valuable feedback on their progression, but also equip teachers and program boards with high-end data for pedagogical and course development purposes. Besides illustrating the developmental methodology, the paper includes reactions and remarks from students on what the tests would provide and what would motivate them to take it.

Keywords: Concept inventory tests (CITE), progression, Courses-concept matrix (CCM), deep learning

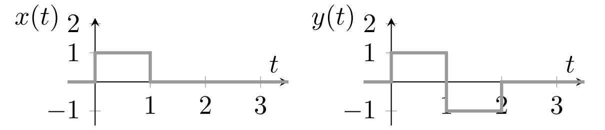
1. INTRODUCTION

For engineering students to successfully complete their degrees and to be able to operate as independent actors in their field, they need to acquire specific procedural and conceptual knowledge. In this paper we focus specifically on the development and assessment of conceptual knowledge, i.e., understanding the underlying principles of core concepts, since previous studies and our own teaching experience have highlighted that these are areas of concern. For example, Surif et al. (2012) conclude from problem solving tasks and interviews that most students are weak in conceptual knowledge. Cracolice et al. (2008) also found that most students employ mechanical algorithms as problem-solving techniques, i.e., memorize the necessary formulas of processes without learning and understanding the concepts. These findings are in line with our own experiences as we have observed that students often struggle to understand fundamental concepts or fail to retain their understanding between courses. An instance which exemplifies this are the answers to the question on the basic concepts on linearity and time-invariance in Figure 1 that was asked in a written exam for third year electrical engineering students in the course *Signals and Systems* at UU. These concepts are considered fundamental in the engineering context and were taught not only in the *Signals and Systems* course, but also in previous courses such as *Transform Methods*. Despite this, only 52.6% (30 of 57) of the students on the course answered the question correctly, a result that made us aware of the lack of attention to the development and consistent assessment of conceptual knowledge throughout our engineering programs.

Assessment in the form of traditional written exams is commonplace in engineering education; unfortunately, a drawback of such exams is that they predominantly assess procedural knowledge, i.e., the ability to complete a specific calculation or task. Limited attention is paid to assessing the students' understanding of fundamental concepts. Moreover, exams are generally designed to assess the goals of specific courses rather than overarching program goals. Finally, the results from individual exams do not provide students with a clear picture of their own progression over time, and it is difficult to use the results from these examinations for pedagogical development purposes since these time-series signals provide limited insights into the students' retention of skills and concepts among

⋆ The research leading to these results has received funding from pedagogical funds at LTU and UU. The authors are listed in alphabetical order. Email-addresses: eva.fjallstrom@ltu.se, steffi.knorn@signal.uu.se, kjell.staffas@angstrom.uu.se and damiano.varagnolo@ltu.se

Consider an LTI system for which to this input signal $x(t)$ corresponds the following output signal $y(t)$:



Derive and carefully sketch the output signals $y_1(t)$ and $y_2(t)$ corresponding respectively to the following input signals $x_1(t)$ and $x_2(t)$, which would be fed into the same system as above:

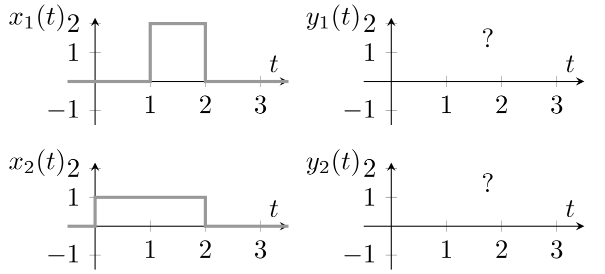


Fig. 1. Example of exam question on linear, time invariant systems from the course *Signals and Systems* at UU, 2018.

courses. From practical perspectives, currently there is no data available that can provide a holistic view on how students develop conceptual knowledge throughout our programs. *The identified need is, hence, for tools for assessing engineering students' acquisition of important concepts throughout their studies that can also be of use for students and teachers plus form a basis for educational development activities.*

To address this need, we are executing a pedagogical development project at our universities that draws on previously developed concept inventory tests (e.g., Savinainen and Scott (2002); Wage et al. (2005)) to develop a test that covers the essential conceptual knowledge for the field of electrical engineering. The aims of our project are 1) to identify a general model for the development of a concept inventory test, 2) to develop a full version of the test, and 3) to identify a suitable format for implementing the test at the electrical engineering program at our two Swedish universities. The long-term goal is to implement the tool so that the extractable data can benefit students, teachers and program development.

This paper does not focus on the concept inventory test that we are developing in itself, but rather on the *methodology* that we have been following for developing this test. In other words, we here present and discuss how we executed the pre-implementation phase of the aforementioned project. The paper is then organized as follows: we discuss the background in Section 2, then propose a methodology for developing a program wide concept inventory test in Section 3. We then review some lessons learned from the feedback received from students at UU and LTU in Section 4, and finally close the paper with a summary, some concluding remarks and future directions in Section 5.

2. BACKGROUND

2.1 The value of conceptual knowledge

Previous studies on the value of conceptual knowledge and the interrelation between conceptual and procedural knowledge indicate that the former forms a solid basis for latter, but also that the latter can be acquired *without* achieving without deeper understanding (see, e.g., Engelbrecht et al. (2012)). As an illustration of this fact, a study about the achievements of physics students (Mazur, 1997, pp. 253) identified a clear pattern where students with a well developed understanding of central concepts also performed well on calculation and application tasks. Some students were able to carry out these tasks successfully on the exams, i.e., demonstrated procedural knowledge without demonstrating a deep conceptual understanding. However, students demonstrating conceptual knowledge rarely failed to demonstrate procedural knowledge.

Other studies also show that students that are educated in conceptual knowledge perform statistically better than students taught for procedural knowledge on both procedural and conceptual tasks (Pesek and Kirschner, 2000; Chappell and Killpatrick, 2003). The value of conceptual knowledge is not generally debated in the context of engineering education since students who have acquired conceptual knowledge tend to be strong on procedural knowledge. Many forms of examination, however, do not explicitly assess conceptual knowledge even if teaching and learning activities might have this focus.

2.2 Concept inventory tests and how they are used

Concept inventory tests can be considered as "diagnostic tests" designed to assess students' understanding of field-specific concepts. Validated concept inventories exist for fields such as, for example, physics, chemistry, statistics, engineering and biology; research is also underway in several other disciplines. The main differences between traditional teacher-authored tests and concept inventories are that the latter focus exclusively on conceptual knowledge, plus that their questions and possible associated answers are subject of extensive research. In other words, when developing a concept inventory test it is essential to: *a)* identify and focus on fundamental concepts, and *b)* formulate questions and suitable answer alternatives that contain distracting answers whose purpose is to expose potential and typical misinterpretations. The form and content of the tests shall also be evaluated so to ensure reliability and validity.

3. A METHODOLOGY FOR DEVELOPING A Concept Inventory Test for Electrical Engineering (CITE)

3.1 The proposed methodology in a nutshell

In this section we propose a methodology to derive a concept inventory for a program such as the CITE being developed at our home universities. While a graphical overview of the whole methodology is shown in Figure 2, its individual parts are described in more detail below.

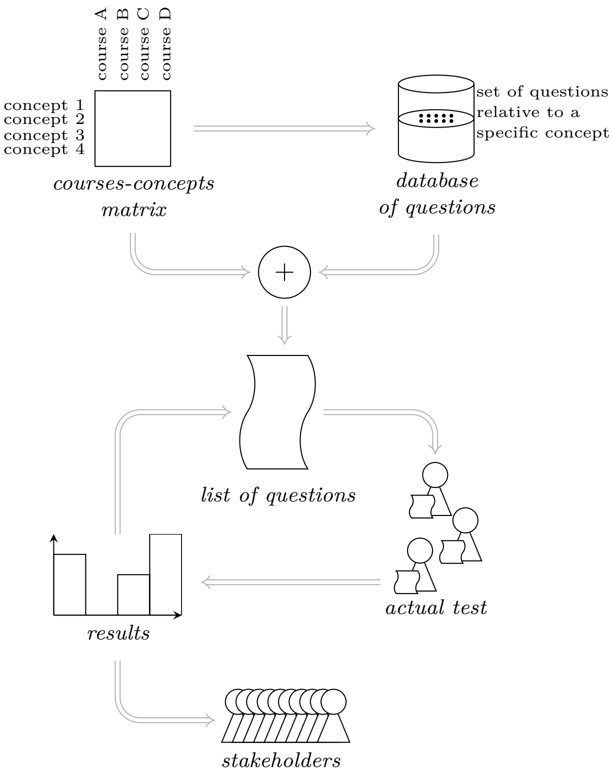


Fig. 2. Graphical overview of the proposed methodology for developing a program wide concept inventory test.

3.2 The Courses-Concepts Matrix

The first step is to determine which concepts should be included in the test. For this, we propose to start by identifying a preliminary list of concepts from the goals and the descriptions of all the courses of the program under consideration and from feedback from the teachers of these courses. After compiling this preliminary list, every teacher shall be asked to rank the importance of the identified concepts for their courses using point scales (e.g., 0 to indicate "not relevant for the course", 1 as "included in the course material", 2 as "significantly important for the course". For instance, if a concept is included in the course goals of a course, 2 points should be selected). The set of these evaluations can be graphically visualized as a table, referred here to as the Courses-Concepts Matrix (CCM) (represented in the top left corner of Figure 2).

Once the CCM has been compiled, the sum of the scores of each concept indicates the relative importance of the various concepts as perceived by the set of teachers. These scores can then be used to rank the various concepts for all courses, and eventually to decide which concepts should be included or not in a CITE. Notice that in this final step the overall importance scale should be also include soft considerations such as trying to include concepts from a reasonable spread of different courses, plus considering whether the potentially includible concepts can easily be tested in the formats described in the following subsections. The decision on how many concepts should be included in the test, moreover, should also depends on practical limitations and constraints such as time needed to complete the test (see also Section 4 for students' opinions about this particular issue).

3.3 Developing a concepts inventory questions database

After having decided on a preliminary selection of concepts that should be included in the test, we propose the process to shift the actual development of the questions for the test.

Since the test should focus on *conceptual knowledge* rather than factual or procedural knowledge (both often assessed in traditional exams), special care should be posed in this crucial step. We suggest to start by inspecting famous concept inventory tests that have been developed in the past (in our electrical engineering concept, two are specially relevant: the Force Concept Inventory (FCI) Hestenes et al. (1992) and the Signals and Systems Concept Inventory (SSCI) Wage et al. (2005)). Moreover, based on the available tests (Hestenes et al., 1992; Wage et al., 2005), it appears that multiple choice questions are the most suitable format for testing conceptual knowledge [1]. We stress that in this step it is not only important to clearly state the question so that it examines a specific concept, but also to carefully choose wrong answering alternatives that, by working as distractors, expose potential and typical misinterpretations. To find those it is thus essential to find and understand common misconceptions. In its turn this is in our experience a difficult task; to solve it we found it beneficial to inspect previous wrong answers to similar questions in different formats (for instance open questions instead of multiple choice questions) plus seek aid from experienced teachers, who typically have a good understanding of what students typically misunderstand.

The questions in a concept inventory should finally not only allow to reveal misconceptions through carefully chosen incorrect answers, but also distinguish misconceptions from the lack of knowledge. In other words, the test should allow to distinguish between students that do not know a concept (lack of knowledge) from students that have misunderstood a concept but are unaware of their misconception. Hence, all questions also include one answer option "I do not know", plus students should be encouraged to select this instead of randomly guessing an answer.

Notice that due to this choice of format (i.e., using multiple choice questions, which is in its turn based on well established tests (Hestenes et al., 1992; Wage et al., 2005)) some concepts might not be suitable to be included in the test. Hence, the selection of concepts based on the ranking in the CCM might have to be adjusted in an iterative fashion.

We finally remark that, in our opinion, another important ingredient for testing conceptual knowledge against factual knowledge is including questions that have never been presented to the students before they actually take the test, i.e., avoid repeating posing the same questions more than once. While this seems to be obvious also in traditional exams, this is particularly important in our context: if students may know questions in advance, then the multiple choice format questions do not allow distinguishing between concepts that have been understood versus answers that have been learned by heart. Since we suggest the test to be taken regularly (see Section 3.4),

[1] Multiple choice questions are also a practical option as they can be graded automatically for large amount of students in short time.

this calls for developing a number of alternative questions for each concept.

3.4 Implementing a CITE

Assume that the CCM has been filled, and that the database of questions in Figure 2 has been populated. Planning to implement a CITE requires then to take several important decisions. The most important one is likely to decide whether participating in the test should be voluntary or compulsory for students. This decision must be based on legal considerations, wider practical consequences (e.g., how to enforce a compulsory test and how to deal with potential re-sits in this case) as well as expected outcomes based on feedback from students, teachers and the program board.

In case the test is to be made voluntary, it appears to be advisable to investigate which options are most suitable to motivate a large number of students to participate in the test. For this, asking students for their opinion on various options seems to us a valuable strategy. Options could include: getting various forms of informative feedback on the their performance (also over time), plus the results being used to improve the teaching and the structure of the program or small benefits in kind.

Various time planning issues have moreover to be considered. First, there is the need to trade-off between short times, that do not demand too much time from the students but prevent assessing a large span of concepts, and long tests, that allow posing a large number of questions and thus gather detailed data. Similarly, there is the need for understanding how often to take the test. The least-demanding option from students perspectives is to take the test once a year; once every term (or, even better, four times a year), on the other hand, would allow to gather more information and form time-series that could be used to infer which concepts tend to be forgotten most often, how fast, and potential correlations / causations among learning / forgetting different concepts (see Section 3.5). Finally, yet another issue is to decide when the test should be taken: the option that most likely would allow teachers to best adjust their teaching based on the test results is to take the test right *before* the start of each teaching year/term/period. However, this time may be occupied with regular (re-)exams – in this case it is reasonable to assume that students would be inclined to prioritize their exams. Instead, tests can also be taken during the first teaching week of the year/term/period, which might inflict less stress onto the students.

To conclude, which option should be chosen as for when implementing the test will thus depend on whether the test will be compulsory or voluntary. Our ansatz is that in case voluntary participation of the students is intended, getting students' feedback in order to find suitable options for them will be essential to motivate participation.

3.5 Using the information collected through a CITE

Assume that a program has been implementing a test like the one here described regularly for all students for some time, so that a large amount of data has been gathered over time. The next crucial step is how to handle and use this data so that it will be useful and significant for all the various stakeholders. At the current stage this step is from our perspective still a *work in progress*. We nonetheless foresee the following potential strategies for providing information to the three main categories involved in our setup:

as for the students, our current guess is that they should receive feedback on their results and progress over time in a suitable, constructive and clear manner. Moreover, to avoid having students focus on learning the questions rather than the concepts, questions should not be handed out; this means that students should in our opinion receive feedback on which concepts were answered correctly or incorrectly, rather than which answers were answered correctly or not. Moreover they should get individual graphical representations of how their scores have been varying in time, and which concepts they should review, specially considering which courses they are supposed to take next;

as for the teachers, they will most likely benefit on getting data on how well students perform on concept questions relative to their course *before* starting to teach. Here we believe that aggregated statistics (e.g., average knowledge levels per concept plus spread of the individual knowledge levels) may be sufficient, but do not exclude that providing individual data per student may be useful. Notice that the latter option requires understanding whether teachers should have access to anonymous data or personal data;

as for the program boards, we expect that they will benefit from anonymous and aggregated test data potentially coupled to meta-information about the courses (e.g., teaching method such as classic or flipped, type of examination, etc.) and changes about the course, in order to make informed decisions on how the program should be adjusted. Notice that here it may be very useful to apply big-data analytics tools on the gathered numerical evidence, so to get holistic indications on potential correlation or causation effects (e.g., teaching a course before or after an other course may influence how well students understand and retain a concept). These findings may be once again useful information for taking evidence-based decisions about restructuring a program.

4. LESSONS LEARNED FROM STUDENT FEEDBACK

In order to understand students' opinion about implementing a CITE in their program, 109 students from the electrical engineering program at UU and LTU were asked to fill out an anonymous questionnaire during autumn 2017. The results that were found to be most insightful and clear are indicated below.

4.1 Expected usefulness of the results

As indicated in Figure 3, students were asked to indicate how useful they expect the test results of a CITE to be for them personally, for other students, for teachers and for the program board on a scale from 1 = "completely useless" to 5 = "extremely useful".

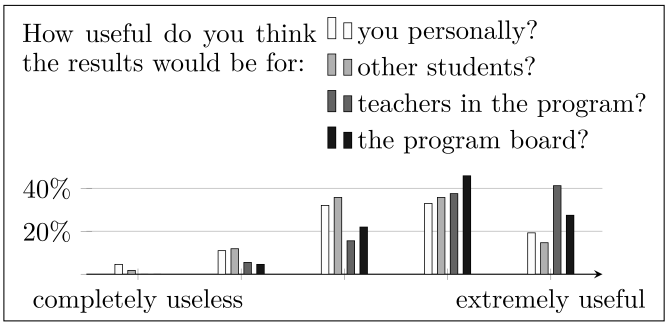


Fig. 3. Anonymous feedback regarding the question how useful they rank the results from CITE for different stakeholder.

As shown in Figure 3, students mostly agree that the results will be interesting for all stakeholders. However, they anticipate the results to be more useful to teachers and the program board rather than for themselves and fellow students. This might be explained by an expectation from the students that the teachers and the program board would take the results into account. In fact, when asked "Which aspects or options would motivate you to participate in taking the test or rather make you less likely to participate or take the test" students indicated increased motivation to take the test if the results would be taken into consideration by the teachers or the program board, as explained in Figure 4.

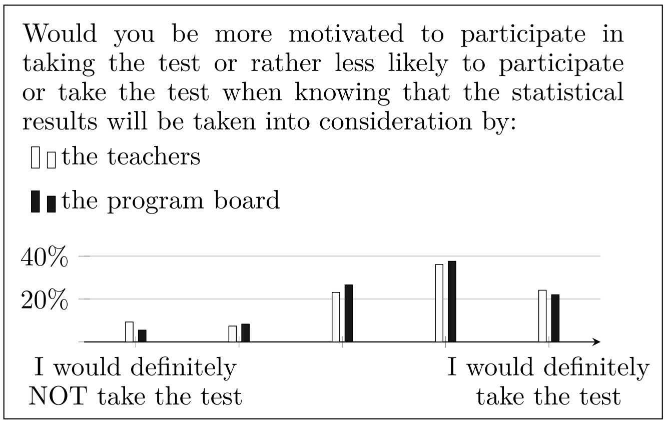


Fig. 4. Anonymous feedback regarding the question if knowing that the teachers or program board would take the results into consideration would motivate them to participate in CITE by taking the test.

4.2 Motivation to participate in the test

Students were also asked to rate other options on whether they would be motivating to take the test or not, again on a scale from 1 = "I would definitely NOT take the test" to 5 = "I would definitely take the test". Out of the presented options, the following were perceived as clearly increasing the motivation of taking the test: "Receive fast feedback on your results", "Have access to your history of results over time", "The result being available only in an anonymous version to the teachers and the program board", "The test being implemented in a convenient setting, at a time that suits you well" and "Getting a fika or coffee or lunch when or after taking the test" [2], as shown in Figure 5.

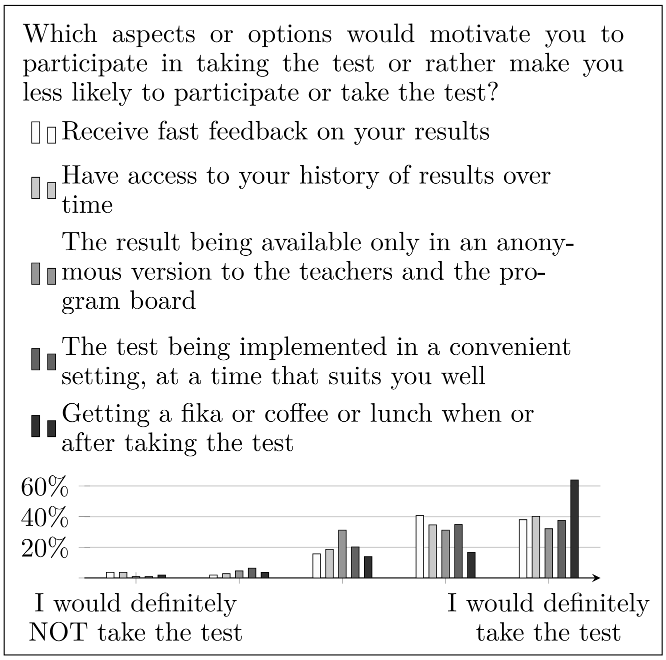


Fig. 5. Anonymous feedback regarding the question what would motivate them to participate in CITE by taking the test.

This in fact matches with the following quotes from some written comments from the students:

- "Fika sounds good"
- "Anonymity is important"
- "This kind of test would in my opinion be a very good addition to the program. I would prefer short and frequent tests. Fika afterwards and the test not affecting grades would help differentiate it from conventional exams."

The second half of the last comment above hints on another important aspect, that would instead of motivating students to take the test, likely have the opposite effect. When students were asked to indicate whether having the results from the test counting towards their academic record, the overall majority said that they would rather not take the test (see Figure 6).

In fact, clearly distinguishing between the traditional exams and a CITE appears to be essential to motivate students to participate, answer honestly, and be open to feedback. Hence, failing to clearly differentiate between exams and a CITE could potentially lead the project to failure. This is also supported by another written comment from a student: "Aren't exams enough tests to show what we learned? I can see the test being voluntary IF one would like to but otherwise not."

4.3 Comparing results with other students

The dislike of students being graded as in exams, or the aversion towards making results from a CITE count

[2] "Fika" is the Swedish word for some small snack, usually coffee or tea accompanied by a bun.

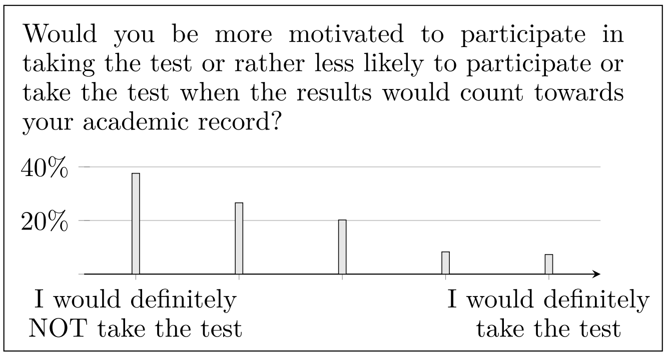


Fig. 6. Anonymous feedback regarding the question if the results counting towards their academic records would motivate them to participate in CITE by taking the test.

towards their academic record might also be connected to the reluctant answers to the question "How important would the result be for you to compare your knowledge to other students" (see Figure 7).

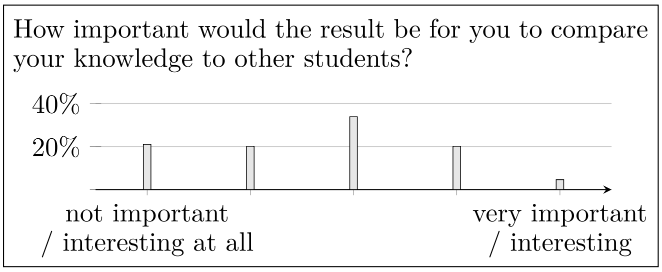


Fig. 7. Anonymous feedback regarding the question how important they find it to compare their test results with other students.

Finally, when asked if "Receive feedback on how you compare to other students" would motivate them to participate in taking the test or rather make you less likely to participate or take the test, about one out of three students answered that then they would rather not or definitely not take the test (see Figure 8).

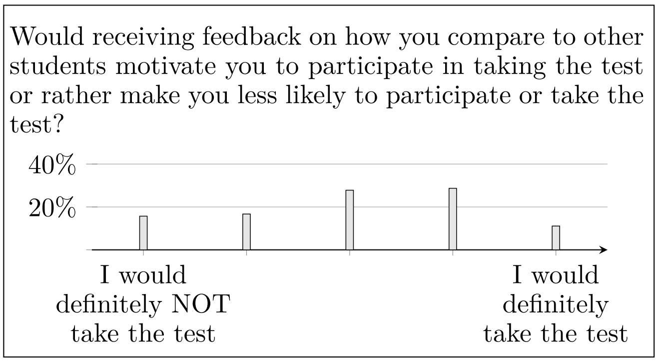


Fig. 8. Anonymous feedback regarding the question whether receiving feedback on how their test results compare with other students would motivate them to participate in CITE by taking the test.

Hence, giving students feedback on how their performance compares to their peers might be leading to less students taking the test. Note that even though almost 40% of the students said that such a comparison would increase their motivation to take the test, it is unclear whether the lack of such comparative feedback would make those students less likely to participate since the opposite question was not asked. One student wrote in the questionnaire: "It's more useful for me to see how I see where I am and what to improve, I don't want to see what other student are and their knowledge. I'm am studying for my gaining not to compare to other students."

5. SUMMARY

It is clear that implementing concept inventory tests on regular basis may provide numerical evidence for better understanding where the students have problems and enact a series of activities / corrective actions that promote deep learning. We foresee that if a significant amount of information becomes available for teachers and program boards, then this will significantly help developing / revamping courses and parts of the program. We finally notice that our pedagogical development project at UU and LTU is currently in a preliminary state where a pilot has been executed; conclusions from the associated results have resulted in the here proposed methodology for developing concept inventory tests. We finally remark that one of the most important aspects to be decided from a program board perspective is whether to make such a test be compulsory or not. This is a keystone to be determined and much of the progress of a project similar to ours depends and decides on this.

REFERENCES

Chappell, K. K. and Killpatrick, K. (2003). Effects of concept-based instruction on students' conceptual understanding and procedural knowledge of calculus. *Problems, Resources, and Issues in mathematics undergraduate studies*, 13(1):17–37.

Cracolice, M. S., Deming, J. C., and Ehlert, B. (2008). Concept learning versus problem solving: a cognitive difference. *Journal of Chemical Education*, 85(6):873.

Engelbrecht, J., Bergsten, C., and Kågesten, O. (2012). Conceptual and procedural approaches to mathematics in the engineering curriculum: Student conceptions and performance. *Journal of Engineering Education*, 101(1):138.

Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *The physics teacher*, 30(3):141–158.

Mazur, E. (1997). *Peer Instruction: A User's Manual*. Prentice Hall, Upper Saddle River.

Pesek, D. D. and Kirschner, D. (2000). Interference of instrumental instruction in subsequent relational learning. *Journal for Research in Mathematics Education*, 31:524–540.

Savinainen, A. and Scott, P. (2002). The force concept inventory: a tool for monitoring student learning. *Physics Education*, 37(1):45.

Surif, J., Ibrahim, N. H., and Mokhtar, M. (2012). Conceptual and procedural knowledge in problem solving. *Procedia-Social and Behavioral Sciences*, 56:416–425.

Wage, K. E., Buck, J. R., Wright, C. H., and Welch, T. B. (2005). The signals and systems concept inventory. *IEEE Transactions on Education*, 48(3):448–461.