# DESIGNING TASKS AND FEEDBACK UTILIZING A COMBINATION OF A DYNAMIC MATHEMATICS SOFTWARE AND A COMPUTER-AIDED ASSESSMENT SYSTEM

Maria Fahlgren, Mats Brunström, Mirela Vinerean and Yosief Wondmagegne

*Karlstad University, Department of Mathematics and Computer Science; maria.fahlgren@kau.se, mats.brunstrom@kau.se, mirela.vinerean@kau.se, yosief.wondmagegne@kau.se*

*This paper reports on the planning of a design-based research (DBR) study, where the main aim is to develop principles in designing technology-enhanced learning environments utilizing a combination of a dynamic mathematics software (DMS) and a computer-aided assessment (CAA) system. The focus is on the design of tasks and automated feedback of high quality so as to enhance first-year engineering students' engagement in and conceptual understanding of mathematical contents. The paper outlines the rationale for the project and highlights theoretical aspects that will be considered in the study. Moreover, some findings from a pilot study that will inform the first cycle of the DBR study are presented.*

*Keywords: Computer-aided assessment, dynamic mathematics software, formative feedback, task design, university mathematics.*

## INTRODUCTION

It is well established that the transition from secondary school mathematics to university mathematics is challenging for many students. The literature highlights several reasons behind this challenge; at university, students meet a new teaching practice, e.g., lecture format (instead of lesson format), larger student groups, less teacher contact, new requirements of learning habits and study organisation (Jablonka et al., 2017). Besides the wide variety in background, interest and prerequisite knowledge among students (Rønning, 2017), many students enter mathematics courses in higher education with insufficient basic mathematical skills (Abdulwahed et al., 2012). This, in turn, leads to unsuccessful study results for many students (Jablonka et al., 2017), which might cause problems, not only in subsequent mathematical courses, but within other applied subjects, e.g., mechanics and electronics, as well (Harris et al., 2015).

To tackle the 'transition problem', many educators in higher mathematics education have introduced continuous assignments to increase students' engagement early during a course, and prevent students from waiting to work on course material until shortly before the final exam (Rønning, 2017). To ensure that students give time to these frequent assignments, they are (most often) graded and constitute part of the course examination. This, in turn, requires a major effort from the teacher in terms of correction work (Rønning, 2017). However, the past decade has seen the rapid development of technology that supports teachers in this time-consuming work by offering automated correction of student responses. A common notion for these types of technology is computer-aided assessment (CAA) systems. Today, many first year mathematics courses in higher education utilize mathematically sophisticated CAA systems, such as STACK and Möbius (e.g., Rasila et al., 2015).

The literature reports several important affordances provided by CAA systems. For example the possibility of randomizing values for variables, parameters and formulas (Rønning, 2017), and the opportunity of providing students immediate feedback on their progress (Rasila et al., 2015), which, in turn, provides support for more independent study by students (Barana et al., 2018). At the same time, researchers in the field of technology-enhanced assessment point out the potential risk of such assessment focusing on lower-order skills of mathematics (Attali & van der Kleij, 2017; Hoogland & Tout, 2018) and solely on the correctness of a final answer (Rønning, 2017) because such types of task and feedback are most straightforward to implement in CAA systems. Consequently, there is scope for designing CAA tasks that address higher-order skills in mathematics as well as for designing feedback that goes beyond categorizing a final answer as being right or wrong (Rønning, 2017).

One possibility to increase the learning potential when using a CAA system is to embed another type of technology: dynamic mathematics software (DMS) (Rasila et al., 2015; Sangwin, 2013). This type of technology is widely recognized as a tool that can promote inquiry and foster students' conceptual understanding in mathematics (Fahlgren & Brunström, 2014; Jaworski & Matthews, 2011). It is the instant feedback on students' action that makes it possible to use a DMS environment as an arena for exploration, conjecturing, verification, and reflection. Even if DMS feedback does not explicitly provide hints on how to proceed, it provides information that could be used in a productive way by the user (Moreno-Armella et al., 2008; Olsson, 2018). However, there is a need for novel types of task to utilize the opportunities provided by DMS environments (Fahlgren & Brunström, 2014; Joubert, 2017).

Although DMS and CAA systems are both in widespread use on their own, there are few studies that have investigated the integration of these two types of technology (Luz & Yerushalmy, 2019). This paper reports on the preparation for a design-based research (DBR) project, which aims to develop principles to guide the design of a technology-enhanced learning environment in which DMS tasks are embedded in a CAA system that (automatically) provides elaborated feedback based on students' responses. It is the cyclic nature of progressive trial and refinement of design principles that makes a DBR approach suitable for this project. Each cycle consists of three main phases: (a) preparation and design, (b) implementation, and (c) analysis and refinement (Bakker, 2018; Cobb et al., 2003). The focus of this paper concerns the first phase, preparation and design, of the first cycle of the planned DBR study. To inform this first phase, a pilot study was conducted in autumn 2020. In the following, we first describe the planned DBR study, including methods for data collection and analysis. Then, we introduce the pilot study and illustrate by an example how the pilot can inform the main DBR study.

## PROJECT DESCRIPTION

In the planned DBR project, the intervention will consist of computer-based mandatory small group activities involving extended task sequences that form part of a calculus course for first-year engineering students (from various programs). The primary outcome of a DBR study is a deeper understanding of *how* and *why* certain instructional interventions work (or do not work), leading to experimentally grounded design principles: in this case, principles to guide the design of a technology-enhanced learning environment in which DMS tasks are embedded in a CAA system that (automatically) provides elaborated feedback (EF) based on students' responses.

### Theories Guiding the Design

In total, the planned study will involve three cycles which will progressively trial and refine the design principles. Each cycle will be guided by a *hypothetical learning trajectory* (HLT), which besides the designed learning activities, includes the intended learning goal of the tasks as well as hypotheses about students' learning processes (Simon, 1995). In the development of the HLTs, including (re)designing of tasks and related feedback, several theoretical perspectives will provide guidance. Since the main focus of the proposed DBR study concerns formative feedback, theories related to different types of feedback will be central, specifically in guiding the development of elaborated feedback provided by the CAA system.

Shute (2008) uses the notion of 'formative feedback' and defines it "…as information communicated to the learner that is intended to modify his or her thinking or behaviour for the purpose of improving learning" (p. 154). Broadly, the feedback information provided to a learner can be of two main types: verification or elaboration (Shute, 2008). The simplest example of verification feedback is whether the student response is correct or incorrect (Narciss, 2008). Verification feedback that also provides the learner with the correct answer to the task is termed 'knowledge of the correct response' (Narciss, 2008). In addition to these types of verification feedback, the literature refers to 'try-again feedback' (Shute, 2008). Elaborated feedback provides the learner with additional information, besides correctness, in various ways. One type of elaborated feedback, suggested by Barana et al. (2018), is to provide hints to guide students towards a solution. In their model of formative automatic assessment in mathematics, they suggest 'interactive feedback' in terms of step-by-step guidance throughout a possible solution process. By asking students to solve simpler tasks, they encourage them to recall previous knowledge and then gradually acquire the knowledge necessary to solve the problem. However, Rønning (2017) argues that there is a risk that this will result in a simpler and less interesting problem. Besides offering conceptual hints or guidance necessary for solving a task, elaborated feedback can provide an explanation for why a particular response is incorrect, or it can consist of a worked-out example (Shute, 2008). Furthermore, the format and timing of feedback presentation can vary. The literature distinguishes between immediate and delayed feedback (Narciss, 2008; Shute, 2008), and according to Vasilyeva et al. (2007), the feedback can be of one or several of the following forms: text, graph, animation, audio, or video. Besides the elaborated feedback provided by the CAA system, the DMS will provide students with feedback based on their interaction with the DMS. This type of feedback is regarded as implicit rather than explicit (Shute, 2008).

Moreover, theoretical aspects related to the design of different types of task utilizing the affordances provided by a CAA system will be important in the DBR study, e.g., example-eliciting tasks (Harel et al., 2020) and other types of task as discussed in the section describing the pilot study (see below). To prompt students to generate examples is not a novel idea – it has been proposed as a way to engage students actively in their development of conceptual mathematical understanding (e.g., Watson & Mason, 2002). Besides these more generic theories, also topic-specific theories will be needed, e.g., learning theories related to functional understanding in mathematics (e.g., Oehrtman et al., 2008).

Altogether, the planned project will imply many important design choices at various levels. To articulate the theoretical rationale for the choices and to analyse them after empirical testing, the design tool of *didactical variables* (Ruthven et al., 2009) will be employed. Put simply, a didactical variable is any aspect of the task (and related feedback), or the task environment, which may influence the unfolding of the expected trajectory of student learning. Next, we will elaborate on the three phases of each DBR cycle:

(a) *Preparation and design.* Except for the first cycle, which will be guided by the pilot study, this phase concerns the revision of the HLT in light of the knowledge gained from the previous cycle(s) and the emerging generic principles. This, in turn, involves (re)designing of the learning activities, i.e. tasks and related elaborated feedback. Crucial in this phase is the identification and articulation of didactical variables attached to the characteristics of the learning activities. Related to these characteristics, hypotheses on student performance, including utilization of the elaborated feedback, are formulated as part of the HLT.

(b) *Implementation.* This is the conduct of the activities, including data collection from students. Mainly, there will be four types of data sources: (i) CAA responses, (ii) surveys, (iii) focus group interviews, and (iv) recordings of student screens. As in the pilot study (described below), the CAA responses will consist of both short (most often individual) answers that will be analysed automatically and group answers to open-ended tasks (e.g., explanation tasks) that need to be analysed manually. In close connection to the implementation of the activities, a survey will be performed to capture students' overall perception, particularly on the feedback provided. To better understand students' perception of various types of feedback (indicated in the survey), we also plan to perform focus group interviews. These will be audio-recorded, and notes will be made to indicate instances related to the HLT (and corresponding didactical variables). However, as van der Kleij and Lipnevich (2020) point out in a recent review "…research provide[s] very limited insights into how student perceptions of feedback relate to engagement with feedback and subsequent meaningful outcomes." (p. 23). Accordingly, to receive information about students' actual utilization of the feedback provided, we plan to collect screen recordings (including audio) from four groups while working on the activities.

(c) *Analysis and refinement.* In this phase, the data analysis process takes place. Data analysis will compare the HLT with the "actual learning trajectory (ALT)" (Bakker, 2018, p. 61), focusing on key didactical variables. Further, it will involve both quantitative and qualitative methods. The preparation for the data analyses will depend on the type of data collected as follows:

(i) *CAA responses.* The CAA system automatically provides descriptive statistics on the degree to which the students have succeeded in performing certain tasks as well as to what extent they have utilized the various types of elaborated feedback provided. The responses to the open-ended questions, on the other hand, need to be prepared manually, as was done in the pilot study.

(ii) *Surveys*. The surveys will primarily consist of closed questions delivered by an online survey tool enabling quantitative data analysis, e.g., descriptive statistics and cross-tabulation.

(iii) *Focus group interviews*. Guided by the notes taken during the focus group interviews, a selection of relevant instances of the audio recordings will be transcribed verbatim. Next, in preparation for a thematic analysis (Braun & Clarke, 2006), the data will be organized into initial codes related to student perceptions of different types of feedback.

(iv) *Screen recordings*. The screen recordings will generate an extensive data set; hence, we need to identify episodes related to the HLT. In these episodes, students' actions will be described and their reasoning will be transcribed verbatim. These episodes will then be organized into initial codes.

Next, in the data analysis process, themes will be generated based on patterns in the initial codes from the screen recordings and interviews (Braun & Clarke, 2006). These themes will then be used to generate conjectures about students' performance as well as their perception and utilization of various elaborated feedback. These conjectures, in turn, could be tested against the other data material (i.e. CAA responses and surveys), looking for confirmation and counter-examples. Altogether, the analysis process will generate the ALT. Finally, the findings (ALT) will be compared to expectations formulated in the HLT. Reasons behind any differences will be discussed within the research team, providing input to the revision of the HLT in the next cycle as well as development and refinement of more generic design principles.

When the three cycles are completed, a retrospective analysis aiming at the finalisation of generic design principles, grounded in their empirical testing in each of the cycles will be made. In contrast to the ongoing analyses (described above), retrospective analysis seeks "…to place the design experiment in a broader theoretical context, thereby framing it as a paradigm case of the more encompassing phenomena specified at the outset" (Cobb et al., 2003, p. 13).

## THE PILOT STUDY

The pilot study involved 256 first-year engineering students taking a first course in calculus. As part of the course, the students were asked to perform two computer-based mandatory small group activities designed for a DMS environment (in this case GeoGebra) embedded in a CAA system (in this case Möbius). The activities involved sequences of various types of task with a focus on the understanding of the function concept. To encourage student collaborations, students were divided into (101) small groups. However, to ensure active involvement by each student, there was a need to embed individual elements. Accordingly, the activities contained both tasks that require a group answer and tasks that require an individual answer.

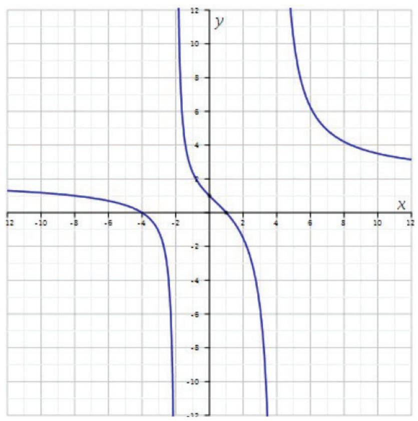
Primarily, the focus was to trial the applicability of different types of task in this 'new' environment as well as to get a deeper understanding of student strategies when performing these tasks. In this way, the pilot provides useful insights into the design of tasks as well as elaborated feedback in the upcoming DBR study. Mainly, three types of task were designed. Firstly, tasks where students were requested to provide examples of functions satisfying specific conditions, i.e. example- eliciting tasks (Harel et al., 2020). In this type of tasks, a design principle was to ask students to provide two examples in order to encourage them to reflect on which parts of the function formula that are possible to vary without affecting the given conditions. Secondly, we constructed tasks where students were asked to determine a function formula for a given graph, e.g., a rational function graph. For both of these types of task, a design principle was to promote students to use the DMS to verify their conjectures before submitting their answer into the CAA system. Finally, we trialed tasks in which exploratory activities in the DMS were central, and where the students were encouraged to explain their empirical findings. In this case, a design principle was to ask students to provide a jointly agreed response to encourage communication and reasoning. Besides the DMS feedback, the CAA system (automatically) provided verificative feedback as well as delayed feedback in terms of worked-out examples illustrating anticipated solution strategies.

The pilot study generated two types of data: student responses to the tasks (generated by the CAA system) and data from an online survey capturing students' overall perception. The findings from the survey indicate that students found the various types of task instructive, and that they found the DMS feedback useful. In contrast, the elaborated feedback in terms of worked-out examples was utilized to a much lesser degree. This finding highlights a need to focus on the development of elaborated feedback that engages students. The data generated by the CAA system offered information about student strategies when performing the tasks, which will provide useful guidance in the (re)design of the tasks and related elaborated feedback. Furthermore, the pilot study provides useful information about methods for data collection and analysis. In the following, we give an example of how findings from the pilot study will inform the first cycle of the main DBR study.

### An Example

The detailed analysis on a sequence of tasks addressing rational functions revealed some unexpected student strategies, i.e. the ALT differed from the HLT. For example, in the task presented in Figure 1, we hypothesized that students should first realize that it must be a rational function with one horizontal and two vertical asymptotes, and then utilize the vertical asymptotes to construct the (factorized) denominator and the horizontal asymptote to conclude that the numerator should be of grade two with the coefficient 2 in front the $x^2$ term. Finally, we expected them to realize that they could utilize the zeros or two other points to finalize the function formula. The analysis of student responses to task ii) (in Figure 1) revealed that almost all students realized that it must be a rational function, and they also utilized the vertical asymptotes to construct the denominator. However, almost half of the students did not utilize the horizontal asymptote as expected. Instead, most of them, utilized the zeros together with one further point, e.g., (0,1) to construct the numerator.

*Below is the graph of the function* **g**.




*i) Use the graph to determine the function formula.*

*Check your suggestion in GeoGebra before submitting it as an answer to the task.*

**Group answer:** $g(x) =$ ____________

*ii) Explain how you used the graph to determine the function formula.*

**Group answer:** ____________________________

**Figure 1. Task as it is presented in Möbius**

This prompted the research team to discuss various options to tackle this particular issue as well as some general principles, both in relation to task design and to the design of elaborated feedback. For example,

- Should tasks be designed so that they cannot be solved without making use of certain key ideas? In the present task, it was the obvious zeros that made it straightforward to find the function formula without using the horizontal asymptote. However, the possibility to use different approaches based on various graph features may promote instructive student discussions.
- Should tasks be designed so that the key ideas are explicit? In this case, it might be an option to indicate the asymptotes in the graph. However, to be able to identify asymptotic behaviour in a graph is a central part of understanding rational functions. Consequently, this kind of scaffolding might simplify the task too much.

Concerning feedback, we discussed the following: when students solve a task without using some key idea, should they then be presented with a question probing that idea, or with a further task that cannot be solved without using that idea, or with a similar task and with feedback asking them to come up with a solution which does use the key idea?

When discussing these options within the research team, both pros and cons were identified. For the task in Figure 1, we decided to develop automated and adapted feedback, which in turn required a redesign of task ii). Instead of asking for an explanation, students were prompted to declare the features (of the graph) used to determine the function formula by choosing among various options (identified in the pilot study). Those students who have not used the horizontal asymptote, were given a new similar task in which they were urged to use the horizontal asymptote.

This example illustrates the complexity of designing tasks and related elaborated feedback. It also shows how information about the ALT could inform the (re)design of tasks to better utilize the affordances provided by a CAA system, i.e. automated correction and adapted feedback.

## REFERENCES


Abdulwahed, M., Jaworski, B., & Crawford, A. (2012). Innovative approaches to teaching mathematics in higher education: A review and critique. *Nordic Studies in Mathematics Education*, *17*(2), 49−68. http://ncm.gu.se/wp-content/uploads/2020/06/17_2_049068_abdulwahed.pdf

Attali, Y., & van der Kleij, F. (2017). Effects of feedback elaboration and feedback timing during computer-based practice in mathematics problem solving. *Computers & Education*, *110*, 154−169. https://doi.org/10.1016/j.compedu.2017.03.012

Bakker, A. (2018). *Design research in education: A practical guide for early career researchers*. Routledge. https://doi.org/10.4324/9780203701010

Barana, A., Conte, A., Fioravera, M., Marchisio, M., & Rabellino, S. (2018). A model of formative automatic assessment and interactive feedback for STEM. In *Proceedings of 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)* (pp. 1016−1025). IEEE. https://doi.org/10.1109/COMPSAC.2018.00178

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77−101. https://doi.org/10.1191/1478088706qp063oa

Cobb, P., Confrey, J., Disessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9−13. https://doi.org/10.3102/0013189x032001009

Fahlgren, M., & Brunström, M. (2014). A model for task design with focus on exploration, explanation, and generalization in a dynamic geometry environment. *Technology, Knowledge and Learning*, *19*(3), 287−315. https://doi.org/10.1007/s10758-014-9213-9

Harel, R., Olsher, S., & Yerushalmy, M. (2020). Designing online formative assessment that promotes students' reasoning processes. In B. Barzel, R. Bebernik, L. Göbel, H. Ruchniewicz, F. Schacht, & D. Thurm (Eds.), *Proceedings of the 14th International Conference on Technology in Mathematics Teaching - ICTMT 14* (pp. 181−188). DuEPublico. https://doi.org/10.17185/duepublico/70762

Harris, D., Black, L., Hernandez-Martinez, P., Pepin, B., & Williams, J. (2015). Mathematics and its value for engineering students: What are the implications for teaching? *International Journal of Mathematical Education in Science and Technology*, *46*(3), 321−336. https://doi.org/10.1080/0020739x.2014.979893

Hoogland, K., & Tout, D. (2018). Computer-based assessment of mathematics into the twenty-first century: Pressures and tensions. *ZDM*, *50*(4), 675−686. https://doi.org/10.1007/s11858-018-0944-2

Jablonka, E., Ashjari, H., & Bergsten, C. (2017). "Much palaver about greater than zero and such stuff"–First year engineering students' recognition of university mathematics. *International Journal of Research in Undergraduate Mathematics Education*, *3*(1), 69−107. https://doi.org/10.1007/s40753-016-0037-y

Jaworski, B., & Matthews, J. (2011). Developing teaching of mathematics to first year engineering students. *Teaching Mathematics and its Applications*, *30*(4), 178−185. https://doi.org/10.1093/teamat/hrr020

Joubert, M. (2017). Revisiting theory for the design of tasks: Special considerations for digital environments. In A. Leung & A. Baccaglini-Frank (Eds.), *Digital technologies in designing mathematics education tasks* (pp. 17−40). Springer. https://doi.org/10.1007/978-3-319-43423-0_2

Luz, Y., & Yerushalmy, M. (2019). Students' conceptions through the lens of a dynamic online geometry assessment platform. *The Journal of Mathematical Behavior*, *54*, 100682. https://doi.org/10.1016/j.jmathb.2018.12.001

Moreno-Armella, L., Hegedus, S. J., & Kaput, J. J. (2008). From static to dynamic mathematics: Historical and representational perspectives. *Educational Studies in Mathematics*, *68*(2), 99−111. https://doi.org/10.1007/s10649-008-9116-6

Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. J. G. Van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (pp. 125−144). Lawrence Erlbaum. https://doi.org/10.4324/9780203880869

Oehrtman, M., Carlson, M., & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' function understanding. In M. P. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics education* (pp. 27−41). Mathematical Association of America. https://doi.org/10.5948/upo9780883859759.004

Olsson, J. (2018). The contribution of reasoning to the utilization of feedback from software when solving mathematical problems. *International Journal of Science and Mathematics Education*, *16*(4), 715−735. https://doi.org/10.1007/s10763-016-9795-x

Rasila, A., Malinen, J., & Tiitu, H. (2015). On automatic assessment and conceptual understanding. *Teaching Mathematics and its Applications*, *34*(3), 149−159. https://doi.org/10.1093/teamat/hrv013

Ruthven, K., Laborde, C., Leach, J., & Tiberghien, A. (2009). Design tools in didactical research: Instrumenting the epistemological and cognitive aspects of the design of teaching sequences. *Educational Researcher*, *38*(5), 329−342. https://doi.org/10.3102/0013189x09338513

Rønning, F. (2017). Influence of computer-aided assessment on ways of working with mathematics. *Teaching Mathematics and its Applications*, *36*(2), 94−107. https://doi.org/10.1093/teamat/hrx001

Sangwin, C. (2013). *Computer aided assessment of mathematics*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199660353.001.0001

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, *78*(1), 153−189. https://doi.org/10.3102/0034654307313795

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, *26*(2), 114−145. https://doi.org/10.2307/749205

Van der Kleij, F. M., & Lipnevich, A. A. (2020). Student perceptions of assessment feedback: A critical scoping review and call for research. *Educational Assessment, Evaluation and Accountability*, 1−29. https://doi.org/10.1007/s11092-020-09331-x

Vasilyeva, E., Puuronen, S., Pechenizkiy, M., & Rasanen, P. (2007). Feedback adaptation in web-based learning systems. *International Journal of Continuing Engineering Education and Life Long Learning*, *17*(4/5), 337−357. https://doi.org/10.1504/ijceell.2007.015046

Watson, A., & Mason, J. (2002). Student-generated examples in the learning of mathematics. *Canadian Journal of Science, Mathematics and Technology Education*, *2*(2), 237−249. https://doi.org/10.1080/14926150209556516