Students mathematical working styles and perception of immediate feedback in a digital assessment system

Siri Ovedal-Hakestad
University of Agder, Norway

The use of digital formative assessment in higher education mathematics is increasing throughout the world. Digital tools come with features such as endless ways of creating tasks and assessing in different ways. However, there is a need of research that focuses on what students do when working with digital tasks and how immediate feedback influence their learning process. This study examines students working styles and immediate feedback in STACK (The System for Teaching and Assessment using a Computer Algebra Kernel). The study reports on interviews with five first-year pre-service teachers who had completed a STACK-test for the first time, in order to get information regarding their working styles and perception of immediate feedback. The results indicate that students have issues regarding mathematical representations in a digital environment and a preference for line-by-line solutions in the feedback.

Keywords: Digital assessment, Feedback, Workings Styles, Syntax

Tools for digital teaching as well as digital assessment have become widespread and accessible in mathematics education (Drijvers, 2018). This was accelerated by the Covid-19 pandemic, and now we have reached a point where computer aided assessment is, in some countries, a part of university mathematics. In addition, previous studies (Faber et al., 2017 (Haelermans et al., 2015; Sung et al., 2016) have shown that using digital learning tools for formative assessment may have positive impact on students' achievement. Drijvers (2020) work show that mathematics students' use of digital tools, in addition to pen and paper, lead to new opportunities and constraints. There can also be room for exploration that may lead to explorative and productive activities, whereas the constraints such as syntactical ones, may lead to frustration (Drijvers, 2020). Therefore, research on students' working styles, where they have opportunities to exploit the opportunities and deal with the constraints a digital tool provides, is essential. The aim of this paper is to explore how students work a digital assessment system and how they approach digital mathematical tasks.

In this paper, working styles is used in line with Weigand and Weller (2001) in their report on changes in students' working styles when they worked with mathematical tasks on computers. Working styles were described in three ways, on tool level, representation level and object level. Their research showed that students worked more experimentally and used more search strategies than expected. However, their study also showed that students did not get a better understanding, but a different understanding of the concept of functions. This study will contain observations of how students use the computer and pen and paper, hence tool level. And how students write mathematics using a keyboard, which is related to representation level.

In the study reported from here, the participants worked with tasks in STACK (the System for Teaching and Assessment using a Computer Algebra Kernel). STACK is designed for formative assessment in mathematics, it is internet-based, and the interactions are through a web-browser (Sangwin, 2013). A distinguishing feature of STACK is the opportunity to create detailed and appropriate feedback for each situation (Sangwin, 2015). In STACK, students enter an algebraic expression into a computer using linear syntax. An option within STACK is validating answers, which means invalid answers are rejected and that students confirm their answer is interpreted correctly by the system before submitting.

This is a qualitative study based on interviews with five pre-service teachers who, before the interview, completed STACK-tests within the mathematical topic of algebra. Students were asked about their working styles and perception of feedback in the interviews, and the aim was to answer the following research questions:

1. What working styles do students in higher education use when working with digital tasks in mathematics?
2. How do students perceive immediate feedback in a digital assessment system?

Conceptual framework

To analyse how students in higher education work with mathematical tasks in STACK and get their view on immediate feedback, the concept of "working styles" was operationalised. The concept of mathematical working styles used by Weigand and Weller (2001) is described in three ways:

> First, it has to be seen on the tool level, e.g., whether one works with tools like paper and pencil, computers, or physical models; second, on the representation level, that describes working with representations like symbols, graphs, tablets, diagrams or pictures, and third, on a mathematical object level, where we see a working style as a sequence of actions that operate on objects and transform these objects. (p. 88)

An object can be an equation, an action can be multiplying, and sometimes one step of this sequence involves several sub-steps (Weigand & Weller, 2001). According to Thies and Weigand (2001), the relationship between object and representation is that the mental objects and their real representations are connected inseparably. In this way, operating with objects on the computer is the same as operating with mathematical objects.

The tool in this study is a computer, which entails special mathematical notations for the objects, and mouse-driven or key commands for different actions. This allows a person to work, for some in new ways, with objects or the representations of objects on the screen (Thies & Weigand, 2001). In this study, the concept of working styles will be used to explain how students work with tasks in STACK, and because of the scope of this paper, only the tool and the representation level are emphasised.

On tool level students had to make their answers using a keyboard. However, they had full access to pencil and paper to write notes. Therefore, it is interesting to observe how they use the tools. On representation level, students work with representations of mathematical symbols on a computer and had to use a keyboard and therefore computational notation to give their answer.

Assessment and feedback

Black and Wiliam (2009) define digital formative assessment as a close view of the teacher, technology, and student interaction within a learning process and their functions. Formative assessment may positively impact learning, and teachers, peers, and students are all actors in formative assessment (Dalby & Swan, 2019). Formative feedback is a key element of formative assessment because it provides the learner with information about their current state of knowledge and their learning goals (Narciss et al., 2014). This research conceptualises formative feedback in line with the definition by Black and Wiliam (1998) which "refers to all those activities undertaken by teachers- and by their students in assessing themselves- that provide information to be used as feedback to modify teaching and learning activities" (p. 82).

Digital learning tools might be an essential aspect of formative assessment in higher education. Technology can be used to improve the quality of the feedback received by students during their formative assessment, especially when considering the following three factors according to (Cosi et al., 2020). First, while using a technological tool, the time factor is immediate. Second, students can receive feedback in different formats such as text, video, audio etc., which enrich the message. And third, facilitation of a range of complementary resources is available on the internet.

According to Kulhavy and Stock (1989), the content of the feedback message should comprise two separate components, verification and elaboration. Verification is a judgement of whether the initial response is right or wrong, while the elaborated component consists of more information than just right or wrong. In this study, there were two different STACK-tests. The first one contained a verification component only, which indicated to the students if their response was right or wrong. The second STACK-test contained both a verification and an elaboration component, where the students also got information about the right response.

Methods

Data collection

The participants in this study were first-year student teachers for grades 5–10 at a university in Norway. All participants had deliberately chosen mathematics as their future subject to teach and the data were collected during their first mathematics course. The first lecture, I informed about the project and offered all students to participate in a two-hour working session every week, where they could choose to work with weekly tasks provided by their lecturer or digital tests created in STACK. The students were reminded of this working session each week. The STACK-tests were in addition to the students' regular work and was created as a supplement to their work towards the final exam. The topic of the STACK-test was algebra, which was part of the course content.

In total, 30 students completed the STACK tests, 13 students agreed to participate in the study, and 5 students participated in an in-depth interview. Early in the semester, three students completed the first STACK test and volunteered for interviews. Later in the semester, two more students volunteered after completing the second STACK test. The participants are, in this paper, named Oliver, Roger, Daniel, Harry, and Jack. It is accidental that all participants are male, because in class there was approximately fifty percent female and fifty percent male. Daniel, Harry, and Jack conducted individual interviews and completed the first STACK test, while Oliver and Roger participated in a group interview and completed the second STACK test.

The STACK-tests

STACK uses the computer algebra system Maxima to interpret students typed algebraic expressions and assign outcomes such as feedback and marks (Sangwin, 2019). STACK is equipped with algebraic equivalence, which makes multiple solutions, as long as they are equivalent, correct. STACK has a built-in interpretation of the students' answers; this means that the students' writing is shown interpreted mathematically before the students give their final answers. Regarding students input in STACK, there are six different options for inserting stars (*). In both tests the option of don't insert starts automatically was selected. This means that if there are any pattern identified the result will be an invalid expression.

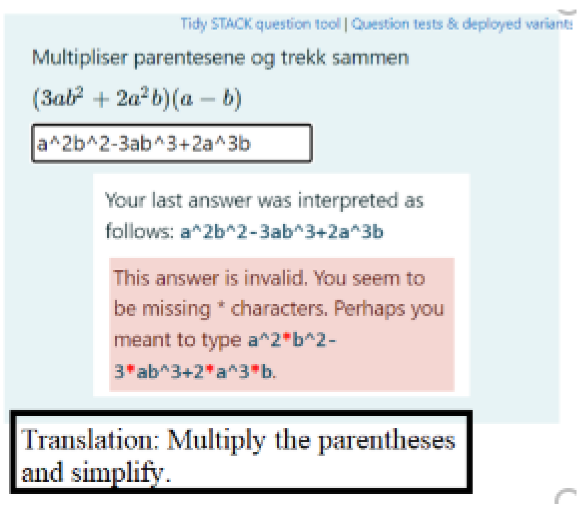

Figure 1. A picture of question 3 which show STACKs interpretations of an invalid expression.

The two tests had 9 similar tasks, and the second test consisted of 6 additional tasks. The first task in both tests concerned acceptance of participation in this study, and students had to actively decide whether to engage or not in this study. The difference between the two tests was primarily the feedback, where the first test had feedback as verification, and the second test had feedback including verification and elaboration. In addition to a difference in verification where the first test a verification was given after each reply, but in the second test, the students didn't get the result until all tasks were answered. All tasks were in the mathematical field of algebra, which was one of the students' mathematical topics this semester.

Because of space limitations and the relevance of this study, not all tasks will be recalled. The aim is to look at what working styles students had in this digital environment and how they wrote the answers into STACK.

Data analysis

To be able to answer the research questions, the interviews were transcribed in their entirety. Afterwards, the transcripts were coded in a process inspired by grounded theory. First, the transcripts were broken into discrete parts, closely examined, and compared for similarities and differences, which Saldaña (2021) refers to as the initial coding process. The transcripts were read statement by statement, and notes were taken on the side. This process created a starting point, which led to further exploration. Some of the notes after this initial process were: improvements, more organized approach to tasks, learning outcome, mastery, multiplication sign, step-by-step response, notation of tasks, and earlier experience.

After preliminary analysis codes that was overarching and related were created (Swanborn, 2018). This led me to two categories which will be elaborated on in this paper: working styles and feedback.

Results

First, I present the findings related to the first research question, addressing working styles students use when solving digital mathematical tasks. Second, findings related to feedback will be presented. Both parts will mainly consist of reproduction of the five interviews with Oliver, Roger, Harry, Daniel and Jack.

Working styles

In this section, ways of working with mathematical tasks digitally are reported. Oliver says that he always writes down each task on his tablet (similar to pen and paper) and calculates line by line. He had a similar approach to the STACK tasks, where he wrote line-by-line what he did in each task. This is a working style he always uses when approaching a mathematical task involving algebra. Oliver says that this way of approaching tasks give him a good overview of what needs to be done in order to solve the task. Roger explains that he tried to calculate in his mind first because he saw some of the tasks to be straightforward. If there were tasks, he could not calculate in his mind, he wrote down some notations to get back on the right track. These notations contained steps he knew how to solve but could not calculate it in his head. He explains that he enjoyed especially a task about quadratic identities, because he immediately saw that this was the case. In that task he had to understand how to calculate with parentheses, which he knew how to do, it became a practice on routine. Roger used STACK to check if his calculations were correct. Harry perceives algebra as a topic he manages well and explains that he did most of the calculations in his mind. He used STACK to guess and check his if the calculations done in his head was correct. If he did not know the answer, he guessed something, in order to look at the right answer after completing the task. He emphasized that since this was an assessment with no impact on his grade, the guessing and checking approach worked fine. If the test had an impact in his grade he would use a different approach, where he would have written down each task and made line-by-line calculations. He had high expectations of his work, especially work that had an impact on his grade. Daniel also explains that since this was a digital test with no impact on his grade, and the level of the tasks was not that high, he did most of the calculations in his mind. He used pen and paper to do small calculations he could not complete in his head. He reports that in this test, and digital tests in general, he does more calculations in his mind and uses the assessment system as a way to check if his answers are correct. Oliver used his tablet, Roger and Daniel used pen and paper in addition to the computer as tools. Harry did use any tool in addition to write notation in this test.

Using STACK for the first time requires writing mathematics with the use of a keyboard. All five students made comments regarding the multiplication (*) sign. The tests were programmed with an option where the program did not insert * characters automatically. Daniel said that remembering to write this sign everywhere because I know it is there. And since we do not usually write it, he forgot to write it in each task. In the interview with Oliver and Roger, a discussion about the multiplication sign occurred because Roger emphasized the importance of knowing that $4a$ actually is 4 times a. However, Oliver said that since he already knows that the sign is there, it is a waste of time to write it every time. The had different views on the importance of this * sign. Oliver was clearly frustrated about this sign, while Roger saw it as a way of writing mathematics digitally. Harry faced a challenge when he was going to write the power of a number. He did not know which symbol belonged to the right mathematical meaning. After spending some time thinking about it, he remembered using GeoGebra in upper secondary school and remembered what key it was. Jack also had some challenges with finding the power of a number. His approach was to write the wrong answer and then look at the right answer in the feedback. Then he could use this to remember which key it was if another similar task appeared. Another challenge Harry reported was that in one task, he wrote the wrong mathematical symbol and got the wrong answer. He said that he did not remember what this symbol was and that this mistake would never happen if he had written these tasks on pen and paper. This shows that students had different approach to finding the right symbol to a mathematical meaning. It is clear that this challenge would not happen if this test was created on paper, and the students were required to write the answers on paper. They knew what the mathematical meaning was, and how to write it on paper, but not the transition to a computer keyboard was difficult. If this test would have had an impact on their grade, Jack would have gotten the wrong answer if he deliberately answered wrong to look at the symbol. It is worth knowing that this can be challenging, because students mark may not reflect their knowledge.

Feedback

Feedback was given differently in the two tests. The first test had a verification component, and feedback was given immediately after each task. In the second test, feedback contained a verification component and details about the correct answer and was given after all tasks were answered. Jack, Roger, and Oliver explained that they would like a step-by-step solution after completing each task. They all agreed that this would help them in the learning process. Oliver and Roger emphasized that it will be easier to know if you are on the right path if you get a response after each task. Oliver said that, in this case, if you have to go through 15 tasks without any feedback, you had no way of knowing you were on the right track. However, if you got feedback earlier, you would know that you were doing right. Jack explains that in some of the tasks he was close to the correct solution, and if he got a line-by-line solution as feedback, he would quickly understand his mistakes. If a similar task would appear, then he could use the feedback in a different situation. He also explains that tasks he did not know how to approach, just by viewing the right answer in the feedback, he would not know how to do it in a similar situation. Daniel elaborated that a line-by-line solution given as feedback would help him figure out exactly where his mistakes were. And that a right answer and a line-by-line solution would help his mathematical understanding and development.

Discussion

This work sheds light on students' own report on working styles and how they perceive immediate feedback in a digital environment. In the case of working styles, the students consider writing mathematics on a computer to be challenging. The students had to use the computer as a tool to communicate mathematics. They also had access to paper and pencil or tablet to write notes during the tests. On tool level, Roger, Daniel, and Harry compared their work done in STACK with mathematical work they usually do on paper. The three of them agreed that if this test were on paper, they would write down more detailed line-by-line calculations on paper. As Usiskin (2004) points out, we have turned to paper and pencil for the last 400 years regarding transformational activities in algebra and arithmetic. Daniel, Harry, and Roger reported that they did most of the calculations in their mind. There are some explanations why this might be the case. First of all, this test had no impact on the students' grades. There might be no incentive for the students to complete all calculations as thorough as they would in a grading situation. Second, there was an opportunity to guess the answers and look at the feedback to see if they were on the right track.

There were some issues on the representation level because the notation used on a keyboard is sometimes different compared to the notation using pencil and paper or a tablet. Some students reported their frustration regarding finding right mathematical meaning, and some experience frustration regarding the multiplication sign, however, using the computer as a tool seemed like something the students are used to in different settings. Harry, Jack, and Daniel report that they have trouble finding the right key on the keyboard that belongs to the right mathematical meaning. The keyboard is the basic interaction with a computer, and according to Sangwin (2013), assigning definite meanings to individual symbols is a problem that needs to be solved.

Since this is the students first meeting with STACK in a test, and with no knowledge about the student's prior experience with writing mathematics digitally, these results were expected.

Students in this study also reported that they forgot to write the multiplication sign in places in line with the regular syntax of mathematics. There is an option in STACK to remove this function; however, as students become teachers, it is fundamentally important to know that the multiplication sign is there. While this issue frustrated Oliver and Roger, Jack saw it as a reminder of the difference between number and variable. This type of syntax error is not interpreted as a wrong result in STACK, but purely a technical error (Sangwin & Ramsden, 2007). A question to be raised for further use of STACK is whether to have this function switched on or off because it might lead to less frustrated students if you remove this function. On the other hand, this function is important, especially for pre-service teachers.

Oliver and Roger completed the second STACK test, where feedback was given after completing all tasks, while Jack, Daniel, and Harry completed the first STACK test where feedback was given after each task. Jack, Daniel, and Roger reported that they would value a step-by-step solution after each task. They all reported they would value more information than just verification or justification and the right answer. Kluger and DeNisi (1996) found that task-level feedback often emphasizes one solution, or strategy, and might therefore be a hinder for the transition to more complex tasks. If a solution to a task is presented to the students, they might lock their future work to only this way of solving a similar task and may therefore not help the students to think about other ways of solving a task. This issue needs more research because what the students want and what they will learn most from is debated.

A report from Shute (2008) says that step-by-step feedback is often more effective than presenting all feedback together. Given a summary after completing the whole test might be perceived as overwhelming, and as a result, students may reject all feedback given. This is in line with what the students in this study report.

Conclusion

Typographical errors can easily be missed by students, resulting in wrong answers. This might lead to frustrated students and should be avoided if possible. There are simple things lecturers can do to avert this problem when introducing STACK to students. When STACK as an assessment system is implemented, students should be taught the usual pitfalls, such as the multiplication sign (if this function is on) and learn which key to push for the most common mathematical operations and meanings. Spending time in class on mathematical representations used in a mathematical digital environment may help the students from spending time on finding the right key for the right mathematical symbol.

Students also report that they would appreciate a line-by-line solution as a part of their feedback. However, previous research is not clear that this is the best for their learning outcome. This study lacks data to conclude about students' learning outcome regarding feedback.

In this study, several students report that they calculated without using written notation. This also raises questions for further investigations. Is this a common case for students when they use computer as a tool to write mathematics? Is this a representations level case? Is this a result only in this study, or are the students working in a digital environment more into guessing and checking? More work needs to be done in order to answer these questions.

References

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, *5*(1), 7-74.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, *21*(1), 5-31.

Cosi, A., Voltas, N., Lázaro-Cantabrana, J. L., Morales, P., Calvo, M., Molina, S., & Quiroga, M. Á. (2020). Formative assessment at university through digital technology tools. *Profesorado, revista de currículum y formación del profesorado*, *24*(1), 164-183.

Dalby, D., & Swan, M. (2019). Using digital technology to enhance formative assessment in mathematics classrooms. *British journal of educational technology*, *50*(2), 832-845.

Drijvers, P. (2018). Digital assessment of mathematics: Opportunities, issues and criteria. *Mesure et évaluation en éducation*, *41*(1), 41-66.

Haelermans, C., Ghysels, J., & Prince, F. (2015). A dataset of three educational technology experiments on differentiation, formative testing and feedback. *British journal of educational technology*, *46*(5), 1102-1108.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin*, *119*(2), 254.

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational psychology review*, *1*(4), 279-308.

Narciss, S., Sosnovsky, S., Schnaubert, L., Andrès, E., Eichelmann, A., Goguadze, G., & Melis, E. (2014). Exploring feedback and student characteristics relevant for personalizing feedback strategies. *Computers & Education*, *71*, 56-76.

Saldaña, J. (2021). The coding manual for qualitative researchers. *The coding manual for qualitative researchers*, 1-440.

Sangwin, C. (2013). *Computer aided assessment of mathematics*. OUP Oxford.

Sangwin, C. (2015). Who uses STACK. *A report on the use of the STACK CAA system, Loughborough University, UK*. https://doi.org/https://hdl.handle.net/2134/18540.

Sangwin, C. (2019). Developing and evaluating an online linear algebra examination for university mathematics. Eleventh Congress of the European Society for Research in Mathematics Education,

Sangwin, C. J., & Ramsden, P. (2007). Linear syntax for communicating elementary mathematics. *Journal of Symbolic Computation*, *42*(9), 920-934.

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, *78*(1), 153-189.

Sung, Y.-T., Chang, K.-E., & Liu, T.-C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education*, *94*, 252-275.

Swanborn, P. (2018). What data to collect. *Case Study Research: What, Why and How*, 73-96.

Thies, S., & Weigand, H. G. (2001). Working Styles of Students in a computer-based Environment-Results of a DFG Project. In *Selected papers from the Annual Conferences on Didactics of Mathematics, Ludwigsburg*.

Usiskin, Z. (2004). A significant amount of algebra. *Nieuw Archief Voor Wiskunde*, *5*, 147-152.

Weigand, H.-G., & Weller, H. (2001). Changes of working styles in a computer algebra environment–the case of functions. *International Journal of Computers for Mathematical Learning*, *6*(1), 87-111.