# Combining self-assessment and automatic-assessment – a mixed methods study

Carina Tusche[1], Daniel Thurm[1] and Shai Olsher[2]

[1]University of Siegen, Germany; tusche@mathematik.uni-siegen.de, thurm@mathematik.uni-siegen.de

[2]University of Haifa, Israel; olshers@edu.ac.il

*This paper introduces a digital learning environment featuring a novel assessment module that merges self-assessment and automatic-assessment. This integration is particularly notable given the limited research on combining these two assessment forms. The module aims to leverage the reflective nature of self-assessment with the efficiency and objectivity of automatic-assessment, potentially enhancing learning outcomes. The paper also outlines the initial phase of a mixed-methods study designed to evaluate this combined assessment module. This study will provide a comprehensive analysis of the module's effectiveness compared to other assessment methods. Through this research, we aim to offer valuable insights into the benefits and limitations of integrating self-assessment and automatic-assessment in educational settings.*

*Keywords: Automatic-assessment, self-assessment, technology, linear functions.*

## Theoretical background

Previous studies have shown that digital formative assessment can support student learning, for example through automatic-assessment and feedback on students' solutions (Harel et al., 2022; Olsher & Thurm, 2021). Furthermore, self-assessment is considered important for the development of students' metacognitive skills and the promotion of personal responsibility for their own learning process (Andrade, 2019). However, there is a gap in research that addresses the **combination** of automatic-assessment and self-assessment to support mathematical learning processes (Olsher & Thurm, 2021).

Self-assessment can be conceptualized as a process in which students reflect on the quality of their work, evaluate how well it aligns with established goals or criteria and make modifications accordingly (Andrade et al., 2019). Self-assessment can promote students' metacognitive and self-regulatory processes by encouraging them to evaluate, reflect on and revise their work (Panadero et al., 2017). However, it is important to emphasize that self-assessment carries the risk of students drawing incorrect conclusions about their learning process.

Technology offers various ways to support formative assessment and formative self-assessment, for example with interactive tasks and adaptive real-time feedback (Harel et al., 2022; Olsher & Thurm, 2021, Olsher et al., 2016). As Harel et al. (2022) showed, students working on digital "example-eliciting tasks" (tasks in which students construct examples that illustrate/support their answers to a given problem) can be supported by automatic "attribute isolation elaborated feedback" (AIEF) by providing information on whether specific predefined mathematical characteristics are present in their constructed examples. In this context, Olsher and Thurm (2021) suggested that learner engagement can be further enhanced if they self-assess their work in terms of the predefined characteristics before receiving the AIEF.

## Designing an EET addressing the positional relationships of two linear functions

Based on the concept of Olsher and Thurm (2021), we developed a digital learning setting using a GeoGebra applet. The learning setting aims to explore the relationship between the parameters, the number of intersection points, and the positional relationships of two linear functions.

The task requires students to construct three examples with different positional relationships between two linear functions (see Figure 1, left, for one example). To do this, students can move points on the graphs, use sliders or enter or change the parameters on an algebraic level. In the task, students should also formulate a conjecture about how the parameters, the positional relationships and the number of intersection points of the two linear functions relate to each other (Figure 1, box on the right). An example of a students´ conjecture could be: "Both graphs always intersect if one function has a positive and one has a negative slope".

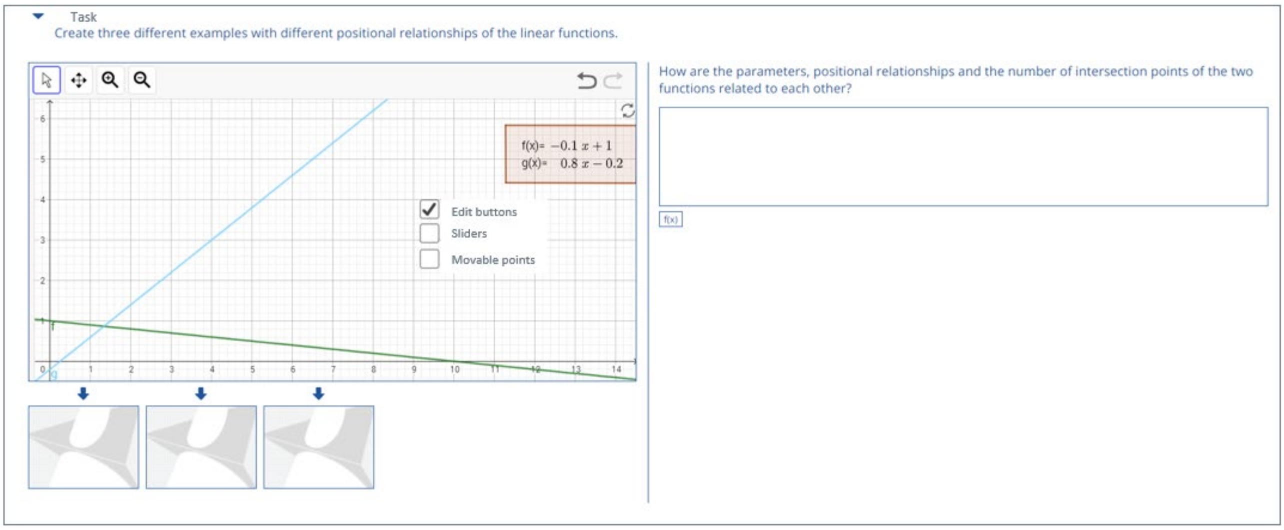


**Figure 1: Task and GeoGebra applet**

After the students have solved the task (Figure 1), they first evaluate each of their constructed examples and decide which of twelve predefined characteristics are present in their constructed examples (Figure 2, left). If they are unsure whether a characteristic is present, they can indicate this by selecting the question mark. Since the students are supposed to explore the relationship between the parameters, the number of intersection points and the positional relationships, the characteristics were constructed in such a way that each characteristic relates to one of the aspects (parameters, number of intersection points, positional relationships). In addition, we defined a characteristic which is not possible to generate with any example to inspire further reasoning processes. After submitting their work, students receive a report consisting of three parts:

**a)** an overview of their self-assessment (can no longer be changed),

**b)** an overview of the automatic-assessment (i.e., overview of which characteristics are present in their examples)

**c)** a combined overview showing conflicts between the self-assessment and the automatic-assessment (Figure 2, right).

Subsequently students work on the task again to improve their task solution.

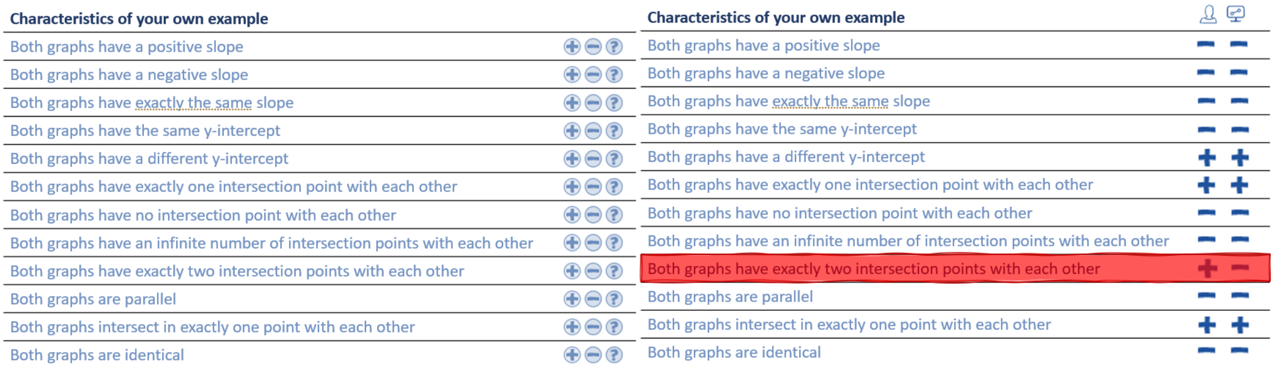


**Figure 2: Left: Self-assessment with predefined characteristics; right: the combined overview with highlighted conflicts**

## Study design

The cluster-randomized mixed-methods study is conducted with approximately 300 ninth grade students divided into three different intervention groups. Within the 45-minute intervention each group engages in the following activities twice:

**A)** *Combination of self- and automatic-assessment:* This group works on the task, then performs a self-assessment with the characteristics (Figure 1, left) and then receives the combined overview highlighting the conflicts between self-assessment and automatic-assessment (Figure 2, right)

**B)** *Only automatic-assessment*: This group works on the task, does not carry out a self-assessment, but receives a report with the automatic-assessment (i.e. which characteristics are present in the examples)

**C)** *Only self-assessment*: This group completes the task and then only carries out the self-assessment with the characteristics, without receiving an automatic-assessment.

To ensure that all intervention groups have the same knowledge at the beginning of the intervention, all three groups receive the same 45-minute introductory phase in which technical terms (e.g. intersection) are learned and repeated. In the introductory phase, no explicit connections are made between the parameters of the linear functions, the number of intersection points and the positional relationships, as these are to be explored in the context of the developed task. At the end of the introductory phase, all students complete a short test to check their knowledge.

## Research goal and methods

This study will investigate the extent to which the different interventions affect

- **i)** the metacognitive activities,
- **ii)** the written conjectures,
- **iii)** the variety of generated examples and

**iv)** the understanding of the relationship between parameters, the number of intersection points and positional relationships of two linear functions.

To reconstruct the metacognitive activities, 4-6 students in each intervention group are filmed, and the recordings are analyzed qualitatively. The written conjectures and the variety of generated examples are reconstructed from the work of the students in the digital learning setting and quantitatively evaluated following the data collection. The understanding of the relationship between parameters, the number of intersection points and positional relationships is determined by a post-test.

## Outlook

It is not expected that one of the three interventions will show the best results for all outcome measures i) - iv). By emphasizing the conflicts (between self-assessment and automatic-assessment, see Figure 2, right), intervention A) could help students to use these conflicts for their own learning process. However, the high cognitive load of the combination of self-assessment and automatic-assessment could also be disadvantageous. For example, it may be that the students concentrate more on the reduction of their conflicts and neglect the revision of their hypotheses. In summary, we expect detailed insights into the mathematical learning processes in different assessment conditions, which can help in the design of digital learning settings that integrate automatic-assessment and self-assessment.

## References


Andrade H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*, 1–13. https://doi.org/10.3389/feduc.2019.00087

Harel, R., Olsher, S., & Yerushalmy, M. (2022). Personal elaborated feedback design in support of students' conjecturing processes. *Research in Mathematics Education*, 1–20. https://doi.org/10.1080/14794802.2022.2137571

Olsher, S., & Thurm, D. (2021). The interplay between digital automatic-assessment and self-assessment. In M. Inprasitha, N. Changsri, & N. Boonsena, (Eds.), *Proceedings of the 44th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, p. 431–440). PME.

Olsher, S., Yerushalmy, M., & Chazan, D. (2016). How might the use of technology in formative assessment support changes in mathematics teaching?. *For the learning of mathematics*, *36*(3), 11–18.

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review, 22*, 74–98.