HOW COMBINING SELF-ASSESSMENT AND AUTOMATIC ASSESSMENT MIGHT HELP TO SUPPORT STUDENT ENGAGEMENT

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

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

[2]*University of Haifa, olshers@edu.haifa.ac.il*

While self-assessment and automatic assessment both carry much potential to enable rich formative assessment there is a lack of research that investigates the combination of both aspects to support student learning. In this paper we present a digital self-assessment module in the context of Example-Eliciting-Tasks using the STEP platform in which learners self-assess their work and subsequently receive automatic feedback about conflicts between their self-assessment and the automatics assessment. We present results from a study that shows that this combination of self- and automatic assessment can stimulate cognitive and metacognitive activities. Finally, we outline the first steps to a more extensive mixed-method study investigating the self-assessment-module with a more extensive set of learners and comparing other forms of assessment.

Keywords: self-assessment, automatic feedback, digital learning environment

INTRODUCTION& THEORETICAL BACKGROUND

Previous studies have shown that digital formative assessment can support student learning, for example, by an automated assessment of learners' solutions (Olsher et al., 2016). Self-assessment is also recognized as important in developing learners' metacognitive skills and promoting ownership of their learning (Andrade, 2019). However, little is known about the combination of automatic and self-assessment (Olsher & Thurm, 2021). The project "Interplay between Self-assessment and Automatic Digital Assessment" (ISAA) (ibid.) aims to address this research gap by investigating the combination of automatic assessment and self-assessment in the context of Example-eliciting-Tasks using the digital formative assessment platform (Olsher et al., 2016)

Formative self-assessment

Self-assessment can be conceptualized as a process in which learners reflect on the quality of their work, assess how well it aligns with stated goals or criteria, and make revisions accordingly (Andrade, 2010). Hence, formative self-assessment emphasizes the role of the learners as independent actors. Self-assessment is considered as one of the five key strategies of formative assessment (Wiliam & Thompson, 2008) and can be implemented through various techniques and mechanisms (Panadero et al., 2016, Andrade et al., 2019). Self-assessment can enhance learners' metacognitive and self-regulatory processes (Ruchniewicz, 2022; Panadero et al., 2017) by allowing them to evaluate, reflect and revise their work. However, it is important to note that self-assessment may be difficult for learners due to high cognitive load for learners. Furthermore, there is the risk that learners may draw incorrect conclusions about their own learning process (Ruchniewicz, 2022). Yet despite the potentials of self-assessment, it is still rarely implemented in classrooms (Kippers et al., 2018).

Potentials of technology for formative self-assessment

Technology has various potentials to support formative assessment and formative self-assessment processes (Olsher et al., 2016; Ruchniewicz, 2022). In addition to interactive sample solutions or dynamic visualizations (Ruchniewicz & Barzel, 2019), interactive tasks and adaptive real-time feedback (Yerushalmy & Olsher, 2020, Olsher et al., 2016) are particularly relevant in the context of this study. One specific example for the potential of digital technology for formative (self-) assessment is the use of digital "Example-eliciting-Tasks" (Nagari-Haddif & Yerushalmy, 2019) which are evaluated in real-time using "attribute isolation elaborated feedback" (AIEF; Harel et al., 2022). For example, in the EET depicted in figure 1 (Harel et al., 2022) learners are asked to formulate a conjecture about which types of product functions can be obtained when multiplying two non-constant linear functions and to construct three examples that support their conjectures using a multiple linked representation (MLR) interactive diagram. Subsequently learners receive AIEF that provides feedback which of 10 predefined characteristics (e.g., the product function has no zero point, the two linear functions have a positive slope) are fulfilled in their examples. Harel et al. (2022) show, that iteratively working with the EET task and the AIEF can support learners in improving their conjectures and to extend the variety of the constructed examples.

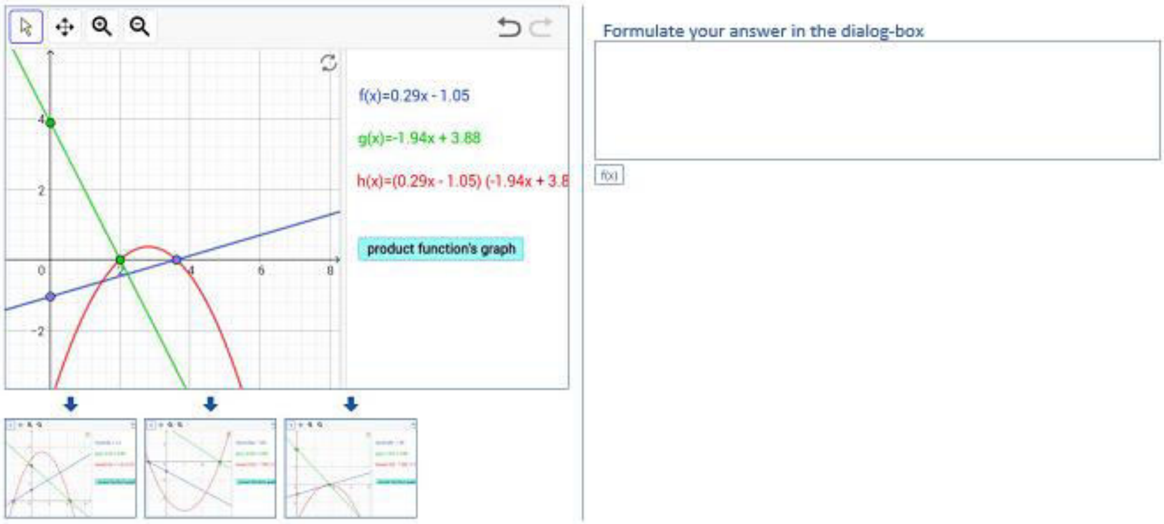

Figure 1. Task

MODULE COMBINING SELF-ASSESSMENT AND AUTOMATIC ASSESSMENT

The present study builds on the work of Olsher & Thurm (2021) who proposed that a combination of self-assessment and automatic assessment may help to spark cognitive conflicts by highlighting conflicts between self-assessment and automatic assessment of predefined characteristics.

Based on this work we conceptualized a self-assessment-module (integrated in the STEP platform) which combines self- and automatic-assessment. After completing the task displayed in figure 1, learners first evaluate which of the predefined characteristics are present in their constructed examples (figure 2). If they are unsure whether a characteristic is present in their example, they can choose the question mark. Subsequently, learners receive a report consisting of three parts: a) an overview of their self-assessment (no longer modifiable, figure 3), b) an overview of the automatic assessment (figure 4), and c) an overview highlighting conflicts between self- and automatic assessment (figure 5).

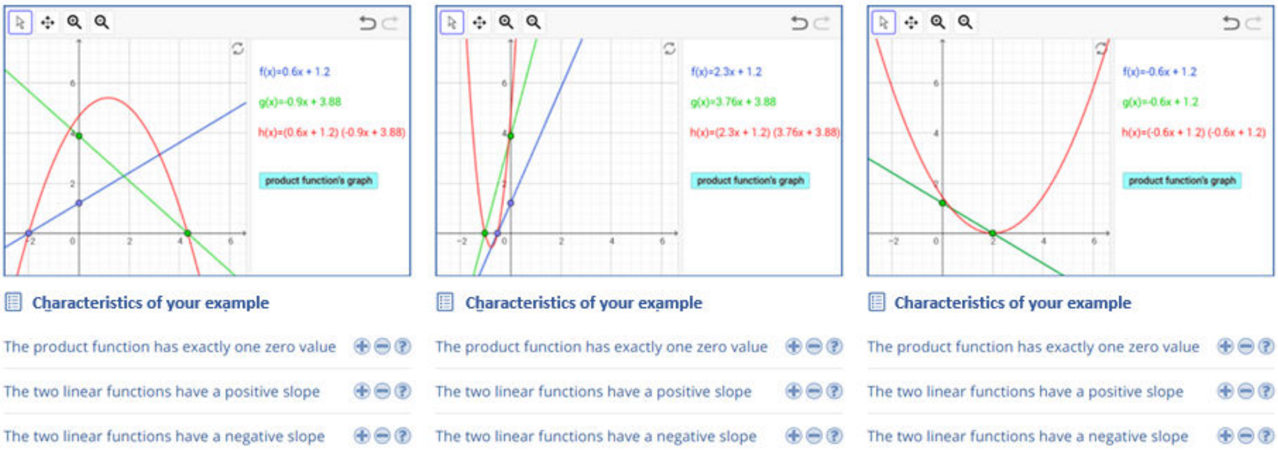

Figure 2. Part of the self-assessment

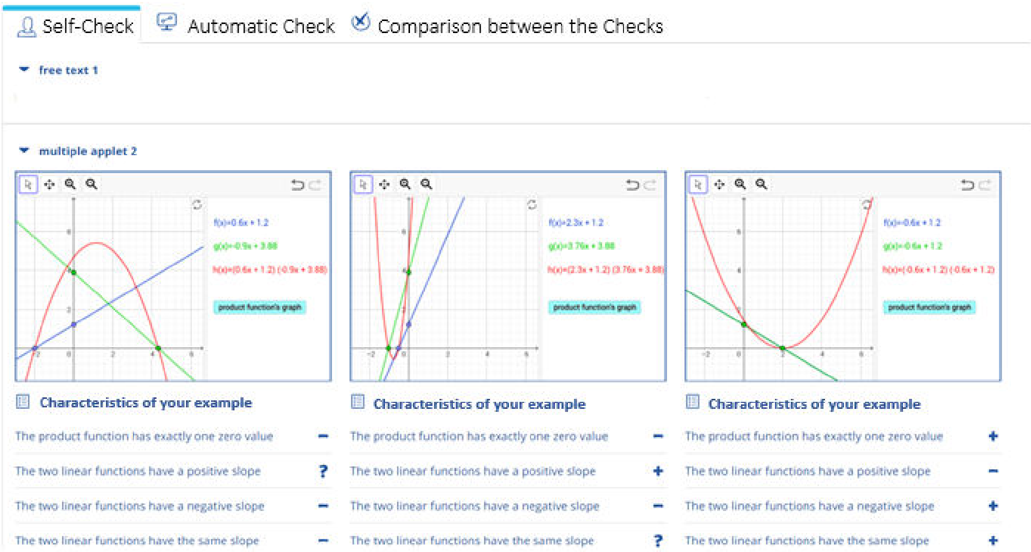

Figure 3. Self-assessment report

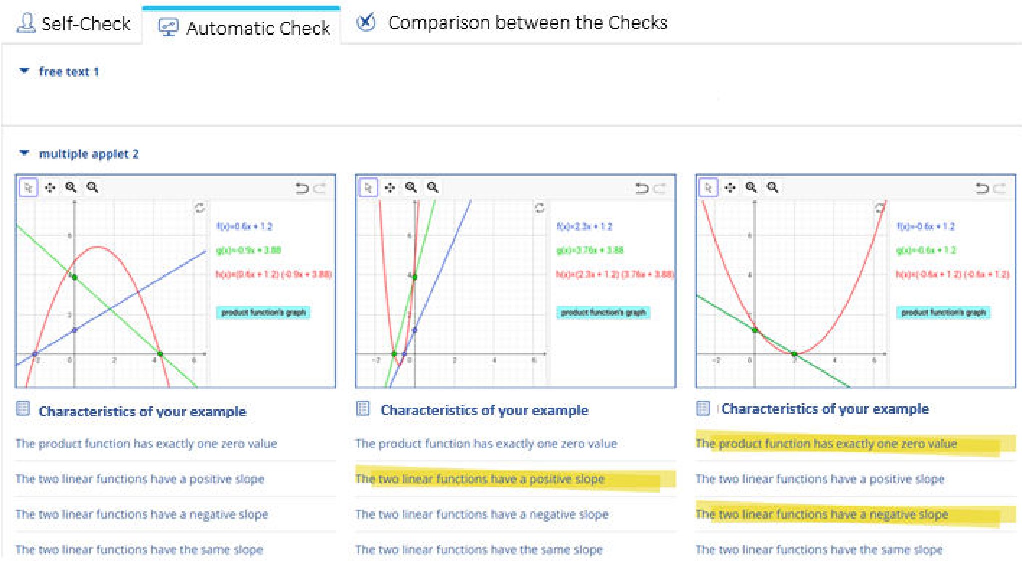

Figure 4. Automatic assessment

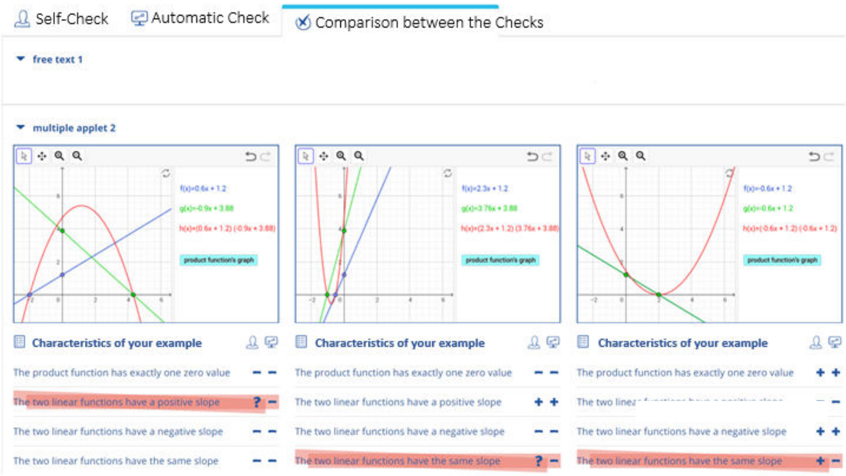

Figure 5. The combined summary

PILOT-STUDY

In a first qualitative pilot video-study with 16 learners (Tusche et al., 2022), we investigated to what extent metacognitive strategies and cognitive activities were sparked by conflicts between self-assessment and automatic assessment. The results showed that most learners were motivated by the feedback given in the report (figure 2-4) to investigate the conflicts between self-assessments and automatic assessment. Moreover, learners had less conflicts (i.e., a more accurate self-assessment) when running through the same EET and the self-assessment module a second time. Hence, the pilot-study suggests that the combination of self-assessment and automatic assessment can encourage learners to be more engaged and active in their formative self-assessment process and that the self-assessment module helps to stimulate cognitive and metacognitive activities.

MAIN-STUDY

In the following step we aim to investigate the self-assessment module with another EET and with a larger set of learners. Furthermore, we aim to compare how learning processes differ between a) learners who work on the EET followed by the *self-assessment module* (figure 2-5), b) learners who work on the EET followed by receiving only the *automatic assessment* (figure 4) and c) learners who work on the EET and subsequently perform *a self-assessment without receiving an automatic assessment*. In the following we will elaborate on the new EET-task that we designed for this endeavour.

Designing an EET addressing the positional relationships of two linear functions

A well-known problem is that learners often have problems with changing representations of functions (Duval, 2006; Schoenfeld et al., 1993). In addition, learners need to understand how the parameters of two linear functions relate to their positional relationship (parallel, identical, crossing) and the number of intersections. This is a complex issue. Figure 6 illustrates the association of these three areas.

Figure 7 shows the EET that we designed to address the aspects. The learners are asked to write a conjecture explaining the association between the parameters of the two linear functions, the positional relationship, and the number of intersections. Because there are three possible positional relationships (parallel, identical, crossing) the learners are asked to create three examples. In addition, learners can manually enter function terms on a symbolic level. Hence learners can change the graph on both the graphical and symbolic level. Dynamic visualization allows learners to easily generate multiple examples and directly observe links between the symbolic and graphical representation.

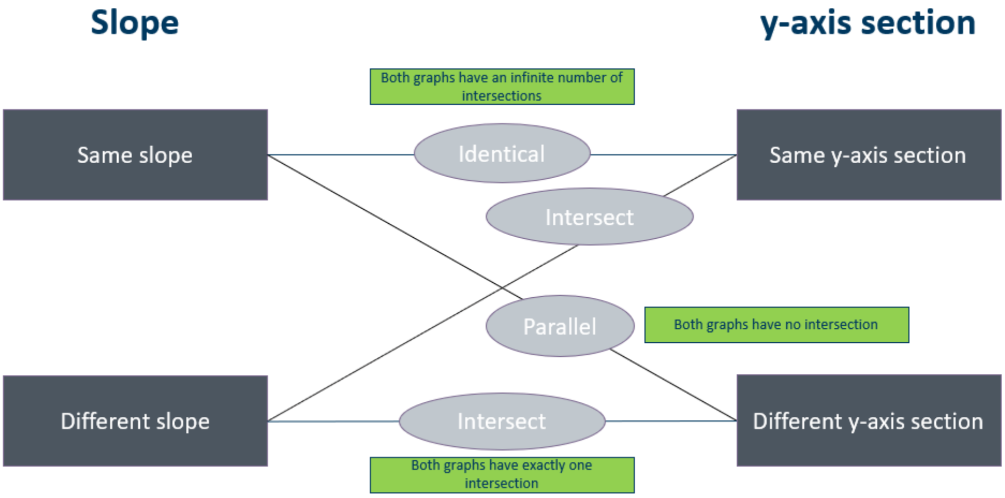

Figure 6. Association of the three areas

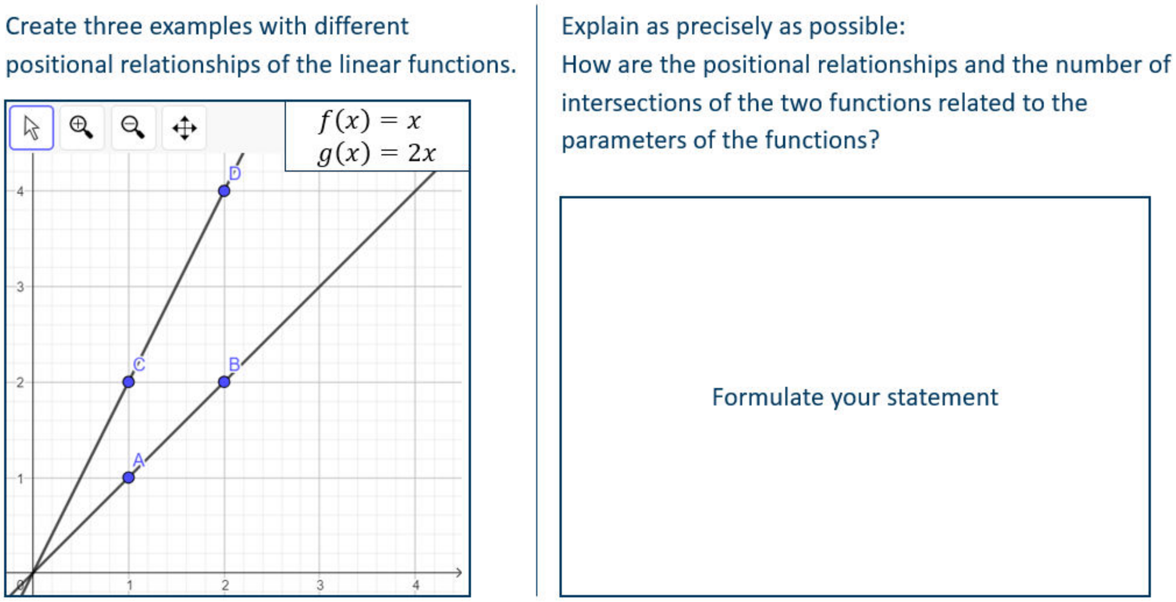

Figure 7. EET adressing the positional relationship of two linear functions

The goal of this task is to be able to establish a relationship between these three areas.

Designing characteristics for the EET

First, attributes were developed that included the three content domains (parameter, positional relationship, and number of intersections). Then we have relied on the distinction made by Harel et al. (2022) who distinguish between three types of characteristics: basic expected characteristics (BEC), integrative expected characteristics (IEC) and challenging characteristics (CC). The first step of this categorization of characteristics was the analysis of the task.

The main goal is to establish the association between the areas parameters, positional relationships and number of intersections. There are two conjectures in this task that can be in the foreground: the number of intersections determines the positional relationship (and vice reverse) the same positional relationships can be achieved with different parameters. Since it is assumed that these three sub-areas will be reviewed and explored in their variability due to the task, the following basic expected characteristics that can support these ideas were first develop:

<u>BEC:</u>

1. Both graphs have a positive
2. Both graphs have a negative slope
3. Both graphs have the same y-axis intercept
4. Both graphs have a different y-axis intercept
5. Both graphs have exactly one intersection
6. Both graphs have no intersection
7. Both graphs have an infinite number of intersections
8. Both graphs are parallel to each other
9. Both graphs intersect at exactly one point
10. Both graphs are identical

Since learners would need to combine other characteristics and their own initial conjectures to formulate the conjecture, the integrative expected characteristic was developed in a further step. Especially the variability of the slope is of importance for the representation of the different position relations. Since the positional relationship "parallel" can only be achieved by an exactly equal slope, this characteristic was supplemented in a further step.

<u>IEC:</u>

11. Both graphs have exactly the same slope

A challenging characteristic was added to activate learners to make new conjectures or check their existing conjectures. This was about the number of intersections, because the learners should deal with the fact that there can be only one intersection or infinitely many intersections of two linear graphs.

<u>CC:</u>

12. Both graphs have exactly two intersections

Word-level formulation has also been an important aspect in composing the character list. An example of a differentiation of the formulation is the characteristic "both graphs have exactly the same slope". This characteristic was chosen to allow learners to make a connection to a parallel positional relationship of the functions. A statement that would also relate to this context would be the statement "both graphs have a different slope". However, this formulation was deliberately not included in the list of characteristics, since it would duplicate the characteristic of exactly the same slope. On the other hand, from the combination with the features "both graphs have a positive/ negative slope" a conclusion can be drawn as to whether the graphs have a different slope or not.

For the learners it is already a high demand to establish the connections between the three areas of characteristic. If they first have to understand the connections between the individual characteristics because they are formulated different, but have a similar meaning, the focus on the actual learning goal can be lost. However, it is important to include both "Both graphs have a negative slope" and

"Both graphs have a positive slope" to represent a contrast. Anyway, the primary focus should be on the learners understanding the connections between the areas of parameters, intersections, and positional relationships. If learners have to continuously create and understand connections within these areas through opposite contexts, the cognitive load for task processing could become too high.

However, it is important to note that it could be a part of the task process that learners may generate examples with different parameters but the same positional relationship. Therefore, during the assessment, they should recognize through the evaluation that their constructed examples are not variable and thus the task has not been solved sufficiently correctly. However, it does not cause any disadvantage for further processing of the task if the learners initially come to such a conclusion. It can be a scaffold for the learners to develop new conjectures about the connections between the different examples. These can then be checked and, at best, improved in a further processing of the task to be able to make more precise statements about the connections.

OUTLOOK

As described, the study pursues the research question of whether the combination of self-assessment and automatic assessment activates learners to engage with their examples and conjectures metacognitively and content-wise more than self-assessment or automatic assessment alone. The study will be conducted with three intervention groups where each group receives the same EET (figure 7) but different forms of assessment (self-assessment module, automatic assessment, only self-assessment). To ensure that all intervention groups have the same knowledge when they start the intervention (e.g. working on the EET) all three groups will receive the same introductory phase in which the learners start with a repetition of already learned technical terms (i.e. parameters, slope, y-axis intercept). Subsequently they are introduced to the different positional relationships of linear functions. In the introductory phase, no explicit connections are made between the different parameters, the number of intersections and positional relationships, since they are to explore this in the context of the developed task. Differences between the intervention groups will be measured by different indicators. First, we will compare the conjectures that the groups formulated. Furthermore, we will investigate the constructed examples. Finally, all intervention groups will do pre- and post-test in which we assess their knowledge with respect to the association between the parameters of the two linear functions, the positional relationship, and the number of intersections. In addition, a subset of learners in each intervention group will be video graphed to capture differences in metacognitive processes. By this mixed-method study, we expect to gain deep insights about potentials and drawback of each of the three types of assessment.

REFERENCES

Andrade, H. L. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In *Handbook of formative assessment* (pp. 90–105). Routledge.

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

Black, P., & Wiliam, D. (2004). The formative purpose: Assessment must first promote learning. *Yearbook of the National Society for the Study of Education*, *103*(2), 20–50.

Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics*, 61(1–2), 103–131. https://doi.org/10.1007/s10649-006-0400-z

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

Kippers, W. B., Wolterinck, C. H., Schildkamp, K., Poortman, C. L., & Visscher, A. J. (2018). Teachers' views on the use of assessment for learning and data-based decision making in classroom practice. *Teaching and teacher education*, *75*, 199–213.

Nagari-Haddif, G., Yerushalmy, M. (2019). *Interactive Example-eliciting Tasks in the Design of Online Assessment in Mathematics* (Doctoral dissertation, University of Haifa (Israel)).

Olsher, S. & Thurm, D. (2021). The interplay between digital automatic-assessment and self-assessment. In Inprasitha, M., Changsri, N. & Boonsena, N. (Eds.), *Proceedings of the 44th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 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. https://doi.org/10.1016/j.edurev.2017.08.004

Panadero, E., Brown, G. L. & Strijbos, J.-W. (2016). The Future of Student Self-Assessment: a Review of Known Unknowns and Potential Directions. *Educational Psychology Review, 28*, 803–830. https://doi.org/10.1007/s10648-015-9350-2

Ruchniewicz, H., Barzel, B. (2019). Technology Supporting Student Self-Assessment in the Field of Functions—A Design-Based Research Study. In Aldon, G., Trgalová, J. (Eds.), *Technology in Mathematics Teaching. Mathematics Education in the Digital Era*, (Vol. 13, pp. 49–74). Springer, Cham. https://doi.org/10.1007/978-3-030-19741-4_3

Ruchniewicz, H. (2022). *Sich selbst diagnostizieren und fördern mit digitalen Medien. Forschungsbasierte Entwicklung eines Tools zum formativen Selbst-Assessment funktionalen Denkens* [Self-diagnosing and promoting with digital media. Research-based development of a tool for formative self-assessment of functional thinking]. Springer. https://doi.org/10.1007/978-3-658-35611-8

Schoenfeld, A. H., Smith, J. P. & Arcavi, A. (1993). Learning: The Microgenetic Analysis of One Student's Evolving Understanding of a Complex Subject Matter Domain. In R. Glaser (Hrsg.), *Advances in instructional Psychology* (S. 55–176). Routledge.

Tusche, C., Thurm, D & Olsher, S. (2022). Investigating the interplay between digital automatic assessment and self-assessment. In C. Fernández, S. Llinares, A. Gutiérrez, & N. Planas (Eds.). *Proceedings of the 45th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 14, p. 305). PME.

Yerushalmy, M. & Olsher, S. (2020). Online assessment of students' reasoning when solving example-eliciting tasks: using conjunction and disjunction to increase the power of examples. *ZDM - Mathematics Education*. *52*, 1033–1049. https://doi.org/10.1007/s11858-020-01134-0