FLIPPED ASSESSMENT IN ENGINEERING MATHEMATICS

Marika Toivola

Satakunta University of Applied Sciences

ABSTRACT

The CDIO framework provides extensive guidelines for improving engineering education quality, yet a critical gap exists in its Standard 11 - Learning Assessment. The framework's treatment of assessment quality enhancement is superficial, concentrating primarily on assessment 'of' learning (AoL) and neglecting the transformative potential of assessment 'for' learning (AfL). While the CDIO syllabus expresses a deep understanding of the importance of self-regulation (González, León, & Sarmiento, 2020) and identity affirmation for success in engineering education, their link to assessment is ignored. The study highlights the necessity for a cohesive integration and reinforcement between AoL and AfL, advocating for self-regulation as a key element of the latter. This is crucial especially in engineering mathematics where serious shortcomings have been identified. To spark a dialogue about the need to update Standard 11, this study presents a practical case from Finnish engineering education, demonstrating how Flipped Assessment (FA) has been specifically developed to facilitate the implementation of Flipped Learning (FL) in teaching of engineering mathematics. The study argues that simply evolving mathematics teaching cultures to align with CDIO standards is not enough; there is a critical need to revolutionize assessment practices as well. In an era where artificial intelligence is challenging conventional assessment paradigms, it is an opportune moment to critically reflect on the ethics of assessment and its validity. As an inherently ethical endeavor, the focus of discussion should shift from the technical validity of assessments to their normative validity. If assessment is detached from its role in nurturing students' mathematical identity and self-regulation, it may lead to engineers who are unprepared for the demands and expectations of their professional careers.

KEYWORDS

Assessment of Learning, Assessment for Learning, Flipped Assessment, Flipped Learning, Engineering Mathematics, Standards: 2, 3, 11

INTRODUCTION

The CDIO (Conceive, Design, Implement, Operate) framework (Malmqvist, Edström, & Rosén, 2020), which is built around a socio-cultural view of learning, offers comprehensive guidelines for enhancing the quality of engineering education, but framework's discussions on improving assessment quality remain superficial. A critical gap exists in CDIO standard 11, Learning Assessment, which, in contrast to the overarching framework, relies on behavioral assumptions regarding learning and assessment. CDIO Standard 11 states the following: "Assessment of student learning is the measure of the extent to which each student achieves specified learning outcomes. Instructors usually conduct this assessment within their respective courses. Effective learning assessment uses a variety of methods matched appropriately to learning outcomes that address disciplinary knowledge, as well as personal and interpersonal skills, and product, process, and system building skills, as described in Standard 2. These methods may include written and oral tests, observations of student performance, rating scales, student reflections, journals, portfolios, and peer and self-assessment."

The standard advocates for the reliability and validity of the methods used to measure the extent to which each student achieves specified learning outcomes (Malmqvist et al., 2020). It acknowledges a spectrum of assessment methods, some of which, like observations and reflections, might be employed formatively. However, the predominant emphasis is on ascertaining terminal learning outcomes rather than utilizing assessment as a facilitator of learning. The prevailing strong psychological measurement paradigm in CDIO Standard 11 does not facilitate the integration of peer and self-assessment methods with self-regulation objectives. The standard predominantly addresses 'assessment of students learning' (AoL), thereby inadvertently marginalizing the dynamic capabilities of assessment 'for' learning (AfL) in nurturing student identity development (Barrow 2006) and enhancing self-regulation capacities (Brown & Harris, 2014). In the CDIO syllabus, the importance of self-regulation is emphasized in sections 2.4.3 (Adaptability, resourcefulness and flexibility), 2.4.6 (Self-awareness, self-reflection, metacognition and knowledge integration), 2.4.7 (Management of time and resources), 2.4.7 (Learning agility, lifelong learning and education), and 2.4.8 (Time and resource management). The strengthening of mathematical identity, which is connected to the fact that the individual sees himself or herself not only as a capable learner of mathematics but also sees mathematics as meaningful to him or her personally, asks especially the form of collaboration listed in sections 3.1.1 (Working in teams), 3.1.2 (Multi-perspective collaboration), 3.2.7 (Inquiry, listening and dialog), 3.2.8 (Negotiation, compromise and conflict resolution), and 3.2.9 (Advocacy).

Serious shortcomings have been identified in the teaching of engineering mathematics (Bennedsen, 2021; Peters & Prince, 2019). Peters and Prince (2019) noticed that engineering students are competent in procedural mathematics but the majority of them have problems in analyzing and resolving a simple engineering problem. They lack the skills to make assumptions, the ability to identify and select appropriate mathematical constructs to create an abstract model, and difficulties in interpreting the results of the model. Obstacles and anxiety in learning mathematics experienced by many engineering students (González et al., 2020) suggest deficiencies in the development of students' mathematical identity. Observed poor self-regulation skills seem to be linked to a lack of teamwork skills and a limited ability to ask questions to solve a problem (Peters & Prince, 2019). To face the problem and strengthen the students' self-regulation, there has been a shift from structured mathematics learning to more autonomous learning, where the learning of mathematics is implicit and not just a straightforward application of previously learned mathematical methods (Peters & Prince, 2019; Treveyan, 2014). What is still missing, however, is a shift from assessment models purely based on judgment to those that actively respond to and support students' self-regulation.

This study responds to the lack of pedagogical debate on the assessment for learning (AfL) and increases the understanding of assessment from a sociocultural theoretical perspective. Flipped Assessment is introduced as a practical example and considered as a purposive cultural intervention for a development of assessment, which is informed and shaped by the values and history of the surrounding society and established school practices. From this perspective, discussing assessment merely as a method (practice or tool) is insufficient; it is also vital to consider its pedagogical rationales in alignment with the goals of CDIO. The CDIO syllabus thoughtfully emphasizes self-regulation as an essential learning objective, suggesting that it should be a fundamental consideration in the design of assessment practices.

ASSESSMENT FOR LEARNING AS A PIVOTAL INFLUENCER OF LEARNING

Although assessment has been recognized as the most important influencer of learning and the summative exams have been identified as the Achilles heel for development of teaching (Rust, O'Donovan, & Price, 2005), assessment practices and research in higher education have still strongly drawn on the psychological measurement paradigm (Boud et al. 2018). The prevailing assessment practices are not dared enough to be questioned by both teachers and researchers (Nieminen, 2021). Changing the assessment approach from judging toward the goals of sustainable development and lifelong learning, which are also valued by the CDIO, is still in its infancy (Topping et al., 2023; Hansen & Sindre, 2023; Bennedsen, 2021).

Assessment is an ethical activity and the kind of discussion that is especially needed is a discussion that moves from the technical validity of the measurement to its normative validity. This involves not only questioning whether we measure what we genuinely value, or whether we measure what is easy to quantify, and finally whether we value what we manage to measure (Biesta, 2009), but also if assessment is detach from its role in nurturing students' mathematical identity and self-regulation, whether this lead to engineers who are unprepared for the demands and expectations of their professional careers. Especially in engineering mathematics, where mathematics as such is not important but mathematical identity development is crucial, teachers should have encouragement to take leaps toward AfL.

There are two cornerstone approaches to assessing student learning: AoF and AfL. AoL refers to the summative measurement of what has been learned, while AfL integrates assessment into the learning process. Although AfL and formative assessment are often considered synonymous, their theoretical basis is different (Baird et al., 2017). These days, the notion formative assessment has become a broad term under which almost any kind of assessment meant to support learning could be categorized (Swaffield, 2011). Initially, it has developed from behaviorist mastery learning theories (Bloom, 1968) and is more related to teachers as the initiators and guides of the process than students as learners (Black & Wiliam, 2009; Baird et al., 2017). AfL, instead, is in line with sociocultural approaches and is more related to students as guides of their own learning (Baird et al., 2017). It is particularly focused on encouraging student self-regulation (Hawe & Dixon, 2017).

To avoid misunderstandings, two other concepts should be distinguished, namely self-regulated learning (Pintrich, 2000; Zimmerman & Schunk, 2011) and self-directed learning (Knowles, 1975). Both involve students' active engagement and goal-directed behavior, but a "self-directed learner controls the learning trajectory as a whole, whereas a self-regulated learner's control is restricted to learning activity" (Cosnefroy & Carré, 2014, p. 4). Thus, the teaching of mathematics within the framework of self-regulated learning encompasses holistic learning objectives rather than solely focusing on mathematics-specific objectives (Toivola, Rajala, & Kumpulainen, 2023). It emphasizes the completion of tasks, rather than directly addressing the objectives of mathematics learning. Consequently, the development of conceptual knowledge in mathematics is contingent upon the utilization of appropriate tools used within self-regulated learning.

Research literature has mainly focused on separation AoL and AfL into mutually exclusive entities based on differences in assessment activities (Taras, 2009; Baird et al., 2017; Bennett, 2011). As a result, two different types of assessment have emerged that do not integrate or support each other. Although AoL and AfL present the different functions of assessment, there is relationship between these two types of assessment and their functions overlap. AoL is always the first part of the AfL process and should not be relieved of all responsibility for supporting learning (Bennett, 2011; Taras, 2009). Likewise, AfL should not be relieved of the teacher's obligation to give grades. Of course, one might ask how justified grades are in engineering mathematics. Whether it be more appropriate than grading to ensure that everyone has sufficient mathematical skills to study engineering? To what extent does grading support or discourage teachers from improving their teaching of mathematics?

Despite efforts to redirect formal testing to reflect the pedagogical underpinnings of assessment (Brown & Harris, 2014), self-assessment within AfL strategies remains in the early stages of development. Self-assessment as a supporter of self-regulation has still received little attention although research suggests that self-assessment could significantly impact self-regulation (Andrade, 2019). In their reviewing literature study Broun and Harris (2013) found three major categories of students' self-assessment practices, namely self-estimation of performance, self-rating, and rubric based judgements. The categories contain procedures like using a model answer as a reference, integrating teacher-evaluation with self-evaluation, self-correction, using a computerized prompt system, self-selected reinforcements or rewards, contributing to the design of a scoring rubric, and judging the accuracy of answers to standardized test items. These all relate primarily to the AoL policies where the focus is on judging the products of students learning.

In general, teachers' discussions about AoL still revolve heavily around the summative exams that are thought to measure whether you have worked hard and studied what the teacher told you to study. Once the exams are positioned to the fundamental method of assessment, discussions about AfL are reduced to discussions of the quantity, timeliness, and effectiveness of feedback in engaging students in desired learning activities. From the socio-cultural perspective on learning, these discussions are limited in value as they are grounded in behaviorist assumptions regarding the nature of learning and its assessment (Shepard, 2005). Because the AfL is not for judging learning, but for learning this study does not follow the steam of research that emphasizes the need for realistic, veridical, or verifiably accurate self-assessment (Butler, 2011). Instead, the study focuses on self-assessment as a process that not only promotes achievement but also empowers students within teacher-driven assessment cultures.

FLIPPED ASSESSMENT AS A LEAP TOWARDS ASSESSMENT FOR LEARNING IN ENGINEERING MATHEMATICS

This section presents the practical implementation of Flipped Assessment (FA) as a pedagogical approach to the AfL in engineering mathematics and considers its theoretical underpinnings to support the development of Flipped Learning (FL). FA is a pre-planned assessment process that is rooted in a socio-cultural learning paradigm, where assessment practices support learning that is created by the learner and their social environment (Toivola, 2020). Throughout the process, students understand that assessment, like learning, is not an effortless process, but requires their active participation. More generally, FA in engineering mathematics supports students' growth as proactive and responsible engineers who understand that things do not happen by themselves but that they have the power to influence them through their own actions.

The context of this article is the Finnish education system, where teachers have high autonomy in both teaching and assessment. Finland has not adopted the internationally popular standards in school test-based accountability policies (Sahlberg, 2007). Initially since 2013, the author has been one of the Finnish FL pioneers at secondary school level (Toivola, Peura, & Humaloja, 2017; Toivola, 2016) and since 2015 developer or FA (Toivola, 2020). From 2021 onwards, the author has been implementing FL and FA in the teaching of engineering mathematics.

The starting point in developing FA in higher education has been that if a teacher has grading responsibility in mathematics, the assessment must be something that produces grades for students. Still, grading should by no means be the most important function of assessment. Grading should rather be considered as a secondary outcome. The main function must be on the pedagogical task of assessment: responsible teaching. In engineering mathematics responsible teaching does not only mean that students master the mathematics to be taught, but also mathematical competences required of engineers. For engineers, it is not mathematics itself that is relevant, but what they use mathematics for is. Companies are looking for engineers who can use mathematics as a tool for success; engineers who have the mathematical competence, the courage to innovate and the courage to fail mathematically in a way that leverages the learning of the whole community.

Initially, an examination of FL is imperative, which has become popular also among teachers in the CDIO network (Leong, Yee, & Kee, 2019; Gommer, Hermsen, & Zwier, 2016). An in-depth understanding of its objectives is essential, as it renders the discourse on FA purposeful and contextually grounded. There are two well-established terms referring to 'flipping', Flipped Learning (FL) and its' precursor Flipped Classroom (FC). FC refers to "an educational technique that consists of two parts: interactive group learning activities in the classroom, and direct computer-based individual instruction outside the classroom" (Bishop & Verleger, 2013, p. 5). FL was launched in 2014 as a response to the prevailing misconceptions among teachers, media, and even researchers about FC, which suggested that flipping was merely a teaching technique without any pedagogical foundation. "Flipped Learning is a pedagogical approach in which direct instruction moves from the group learning space into an individual learning space, and the resulting group space is transformed into a dynamic, interactive learning environment where the educator guides students as they apply concepts and engage creatively in the subject matter" (Flipped Learning Network, 2014).

Both FL and FC have been extensively meta researched as mathematics teaching methods that transform a traditional teaching paradigm (Fung, Besser, & Poon, 2021; Lo, Hew, & Chen, 2017; Naccarato & Karakok, 2015) but there is little research that pays attention to teachers as creators of FL and their pedagogical rationales while implementing FL. In their study Toivola, Rajala, and Kumpulainen (2023) found three main pedagogical rationales for FL in teaching mathematics, namely, individualizing learning, fostering self-regulated learning, and fostering engagement. Individualizing learning emphasizes attempts to differentiate and humanize learning mathematics in heterogeneous student groups. Fostering self-regulated learning highlights the teachers' emphasis on students' responsibility in goal-oriented activity that is supported by self-paced learning. Fostering engagement is related to the teachers' attempts to create a personally motivating learning environment for students.

The author's FL practice in engineering mathematics is strongly based on self-paced learning. In every mathematics course, the units to be learnt are presented in a learning environment, supported with textbooks, exercises, correct answers, and instructional videos, which students go through by themselves at their own rate. Students can adjust their time use for each engineering mathematics course between two months and one academic year. Exam readiness is up to students too; each course has two mandatory FA procedures, with exam days available monthly. A significant structural change here is that the resources received will no longer be earmarked for individual courses. Instead, all resources will be pooled together for students to access, allowing them to use a maximum of 100 hours of teacher's guidance for any single mathematics course. In practice, rather than designating specific times for individual mathematics courses, the schedule features weekly 'mathematics in-class work'. The classroom is open to all students, regardless of their enrolled course, providing a collaborative space for study and guidance.

For a student, FA appears as a two-phase mathematics exam of varying levels, where 'you get a second chance' as you can return the same exam twice. The possibilities to schedule exams and choose the levels of the exams will encourage students to adjust and set their own learning goals, to develop control over their own learning, and to meet competence requirements. In the first 90-minute-long individual phase of the exam, the student's so-called 'failure-level' is determined. During this self-assessment phase, they take ownership of their assessment by setting targets for their desired exam grade. They have the autonomy to choose whether to take the exam at grade levels 1-2, 3-4, or 5, which is the highest grade in Finnish higher education. Every exam of different levels includes 7 tasks, and the failure-level is reached (the exam is passed as the level chosen) if at least 3 out of 7 tasks are correct. In this phase, the steps have been taken toward differentiation and humanization of learning mathematics in heterogeneous student groups. During the second 30-minute collaborative phase, students who have passed the exam work together with one or two peers undertaking the same level exam to improve their overall performance. In this collaborative peer-assessment phase, they collectively review the exam, identify and rectify mistakes, and strive to complete any remaining tasks. This phase is not just used to promote achievement but should be seen as an attempt to promote co-regulation by empowering students during the assessment and by motivating them to become learning resources of each other. The formative use of summative tests is manifested here when the students focus on areas of their weaknesses which they subsequently discuss with their peers. Simply identifying errors is not enough; for a task to be ultimately interpreted as correctly done, students must correct their answers and ensure that the teacher fully understands their corrections. The aim of this activity is not only to increase students understand the assessment process and focus their efforts on improving (Taras, 2009) but also to increase students' positive perceptions of themselves as learners of mathematics and to increase their understanding of the humanistic side of mathematics through their own mistakes (Borasi, 1994; Shepard, 2005). Still, the collaborative phase in FA is not compulsory. Students can continue to complete the exam for an extra 30 minutes on their own if they wish and return the exam only once. Overall, FA fosters a positive learning environment for mathematics, encouraging student engagement and treating mistakes as valuable steps toward self-regulation. It cultivates students' skills in monitoring and guiding their own learning, enhancing their understanding of their roles and responsibilities in making informed choices in their educational journey. The final grading rules are as follows: For exams at levels 1-2 and 3-4, if 3 of 7 tasks are correct, the grade is 1 or 3 respectively. If 5 of 7 tasks are correct, the grade is either 2 or 4. For exam at level 5, if 3 tasks are correct, the grade is 4, and if 5 tasks are correct, the grade is the maximum 5.

FA requires a massive amount of preliminary work. Preparing mathematics exams that cater to various levels of competence is a tedious process for the teacher. During one academic year, the author conducted eight scheduled FA sessions, each of which offered the opportunity to take exams in several different courses. With 21 questions in one exam, this is a huge bank of mathematics tasks required for each FA session. To lighten the workload, artificial intelligence has been utilized to make different exam versions from the same exams. During the period, students completed up to three mathematics courses, which entailed participating in at least six FA processes. The most common number of courses completed ranged from one to two, yet there were also students who did not participate in any exams.

DISCUSSION

This study contributes to novel insights into learning assessment in engineering education, particularly through the lens of mathematics education where serious shortcomings have been identified (Bennedsen, 2021; Peters & Prince, 2019). Although CDIO framework (Malmqvist, Edström, & Rosén, 2020) is built around a socio-cultural view of learning its standard 11, Learning Assessment, relies on behavioral assumptions regarding learning and assessment. While the issue of good engineering education is seen complex, shaped by social interactions and cultural contexts, learning assessment is viewed as a simple process driven by externally controlled reinforcement and punishment. Consequently, Standard 11 falls short in effectively fostering the development of self-regulated engineers, a fundamental aim of the CDIO standards. To meet the gap, the study introduces Flipped Assessment (FA), underpinned by socio-cultural theories of assessment for learning (AfL), as a purpose to broaden the discourse on assessment by providing a tangible, alternative method for evaluating engineering mathematics. In general, initiating critical discussions about assessment practices proves difficult, and there’s a noticeable reluctance to question the status quo in these practices (Nieminen, 2021).

Instead of taking the grades easily given by teachers as the truth, we should cautiously consider not only their legitimacy, but above all what we are assessing, what they are used for and whether they are useful quantities for this purpose (Baird et al., 2017). Although there is evidence of the unreliability of teachers’ grades (Falchikov, 2005; Leach, 2012; Brown et al., 2015; Andrade, 2019) and that final summative exams can be considered as a silent killer of learning (Mazur, 2013), there remains a reluctance to challenge the prevailing assessment of learning (AoL) paradigms. The idea that assessment is a formulaic process, which impartially generates grades, excludes the influence of teachers’ personal opinions, and can be applied effectively at the conclusion of any learning event, is appealing to both teachers and students. Students, having become adept at navigating the terrain of summative exams, know precisely how to prepare for them.

A successful transition to AfL practices requires a clear understanding of the objectives of the new assessment approach and the methods to achieve them. FA prompts critical consideration of the extent of a student’s responsibility in assessment and the ethical implications of seeking peer help to enhance grades. Rather than seeing the assessment responsibility being transferred to the students, FA should be seen as a shared control process (Kirschner and van

Merriënboer, 2013), where the teacher creates exams of varying difficulty, allowing students to choose their level of engagement. To prevent the exams from merely testing the ability to replicate known solutions, each level includes challenges that students are not expected to solve independently. Such tasks support the learning of mathematics in sustained and meaningful ways in collaborative learning situations (Kilpatrick, 2014; Shepard, 2005). Students need support to accept such challenges in assessment situations and to see setbacks as an essential part of their journey toward self-regulation. In FA similarities are sought with Carol Dweck's ideas about growth mindsets, which were told to the students during the introductory lecture. According to Dweck (2006), there are two kinds of mindsets: the growth mindset and the fixed mindset. In engineering, a growth mindset is crucial for innovation and problem-solving. A student with a growth mindset likes the challenges of learning and sees failure as an opportunity to improve. A student with a fixed mindset behavior is regulated by fear of failure, which may prevent the student from even trying and thus hinders the formation of collaborative learning environment. Students need to understand and accept that an exam serves its purpose in the AfL process only when it can identify areas needing improvement. It is a rather fruitless idea to use a measure that indicates that there is nothing to improve. An exam in which a student answers all questions correctly is, from a formative perspective, ineffective in fostering a collaborative learning environment and further advancing the development of self-regulation.

Like Flipped Learning (FL), FA requires students' own ability to take responsibility for their learning. FA may not reach students who struggle with setting deadlines or specific exam dates for their studies. In assessment situations, FA reveals the underachievement for students themselves. It appears that students who have neglected their studies do not gain much from the collaborative phase of FA, as they tend to either skip it entirely or associate with classmates who share their lack of preparation. Conversely, students who are motivated to achieve high grades make effective use of the collaborative phase and seem to enjoy it. For a teacher, it is incredibly rewarding to observe the positive dynamics of group activity that makes the idea of reverting to traditional summative assessments unappealing. Interestingly, not one student has shown interest in returning to traditional summative exams. Instead, the students seem to view FA as a privilege. This suggests that FA genuinely adds value to the educational experience and succeeds in redefining students' perceptions of assessment. In assessment situations, FA can highlight underperformance to students themselves. It seems that students who have neglected their studies benefit little from FA's collaborative phase, as they either skip it entirely or join peers with similar preparation levels. In contrast, students aiming for high grades effectively seem to utilize and enjoy the collaborative phase. For teachers, observing the positive group dynamics makes the thought of returning to traditional summative assessments unattractive. Although not all students have been satisfied with the use of FL as a teaching approach, it is noteworthy that not a single student has expressed a desire to return to traditional exams. After personal experience, students regard participation in FA as a privilege. This suggests that FA succeeds in enhancing the educational experience and changing students' perceptions of assessment. Further research is needed not only on the extent to which FA can support the development of self-regulation and the mathematical identity necessary for engineers, but also on the extent to which it affects students' goals for learning mathematics. The author will begin this empirical study next fall with new engineering students.

FA acts as a deliberate cultural intervention in the evolution of assessment, fitting with existing practices without challenging the traditional grading system. Thus, FA can be integrated into in-course assessments while still accommodating a final exam. In such a context, FA provides teachers with valuable firsthand insights into the effects of AfL on engineering mathematics education and offers fresh perspectives on potential advancements in assessment practices.

Concurrently, it is critically important to consider the pedagogical underpinnings of assessments that align with the CDIO goals. Given the lack of a reliable and valid method for measuring the achievement of CDIO objectives, our focus in CDIO Standard 11 should not be merely on AoF and how accurately we can assess the objectives outlined in the CDIO syllabus, but rather on AfL and how we can leverage assessment to facilitate the achievement of the CDIO framework's objectives.

FINANCIAL SUPPORT ACKNOWLEDGEMENTS

The author received no financial support for this work.

REFERENCES

Andrade, H. L. (2019). A Critical Review of Research on Student Self-Assessment. *Front. Educ.* 4:87.

de Araujo, Z., Otten, S., & Birisci, S. (2017). Mathematics teachers' motivations for, conceptions of, and experiences with flipped instruction. *Teaching and Teacher Education*, 62, 60-70.

Baird, J., Andrich, D., Hopfenbeck, T.N., & Stobart, G. (2017). Assessment and learning: fields apart?, *Assessment in Education: Principles, Policy & Practice*, 24:3, 317-350.

Barrow, M. (2006). "Assessment and Student Transformation: Linking Character and Intellect." Studies in Higher Education 31 (3): 357–72.

Bennedsen, J. (2021). Assessing Students' Professional Criticism Skills – A Mathematics Course Case. Proceedings of the 17th International CDIO Conference, RMUTT and Chulalongkorn, Thailand

Bennett, R. E. (2011). Formative assessment: a critical review. *Assess. Educ*. 18, 5–25.

Biesta, G. (2009). Good education in an age of measurement: on the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability*, 21, 33–46.

Bishop, J. L., & Verleger, M. A. (2013). *The flipped classroom: a survey of the research. In 120th ASEE National Conference and Exposition, Atlanta, GA* (Paper ID 6219). Washington, DC: American Society for Engineering Education.

Black, P. J. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.

Bloom, B. S. (1968). Mastery learning. *Evaluation comment*, 1(2), 1-12, Los Angeles: University of California at Los Angeles, Center for the Study of Evaluation of Instructional Programs

Boud, D., Dawson, P., Bearman, M., Bennett, S., Joughin, G., & Molloy, E. (2018). Reframing assessment research: Through a practice perspective. *Studies in Higher Education*, 43(7), 1107–1118.

Brown, G., Andrade, H., & Chen, F. (2015). Accuracy in student selfassessment: directions and cautions for research. *Assess. Educ*. 22, 444–457.

Borasi, R. (1994). Capitalizing on errors as "springboards for inquiry": A teaching experiment. *Journal for Research in Mathematics Education*, 25(2), 166–208.

Brown, G. T. & Harris, L. R. (2014). "The Future of Self-Assessment in Classroom Practice: Reframing Self-Assessment as a Core Competency." Frontline Learning Research 2 (1): 22–30.

Hansen, G. & Sindre, G. (2023). Assessment and Feedback Across Various Outcomes in Project-Courses: A Department-Wide Study. *Proceedings of the 19th International CDIO Conference*, NTNU, Norway, 913-922.

Harris, L. R., & Brown, G. T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: case studies into teachers' implementation. *Teach. Teach. Educ*. 36, 101–111.

Butler, R. (2011). Are positive illusions about academic competence always adaptive, under all circumstances: New results and future directions. *International Journal of Educational Research*, 50(4), 251-256.

Cosnefroy, L., & Carré, P. (2014). Self-Regulated and self-directed learning: Why don't some neighbours communicate? *International Journal of Self-Directed Learning*, 11(2), 1–12.

Dweck, C. (2006). *Mindset: The new psychology of success*. New York, NY, US: Random House.

Falchikov, N. (2005). *Improving Assessment Through Student Involvement: Practical Solutions for Aiding Learning in Higher and Further Education*. London: Routledge Falmer.

Flipped Learning Network. (2014). The four pillars of F-L-I-P. Retrieved from https://flippedlearning.org/definition-of-flipped-learning/

Fung, C. H., Besser, M., & Poon, K. K. (2021). Systematic Literature Review of Flipped Classroom in Mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, 17(6), em1974.

Gommer, L., Hermsen, E., & Zwier, G. (2016). Flipped Math, Lessons Learned from a Pilot at *Mechanical Engineering. Proceedings of the 12th International CDIO Conference*, Turku UAS, Finland

González, A., León, M. P., & Sarmiento, M. A. (2020). Strategies for the Mathematics Learning in Engineering Cdio Curricula. *Proceedings of the 16 th International CDIO Conference*, Chalmers UT, Sweden, (2), 206-215.

Hawe, E. & Dixon, H. (2017). Assessment for learning: a catalyst for student self-regulation, Assessment & Evaluation in Higher Education, 42:8, 1181-1192.

Kilpatrick, J. (2014). Competency frameworks in mathematics education. In Lerman S. (Ed.), *Encyclopedia of mathematics education* (85–87). Netherlands: Springer.

Kirschner, P. A. & van Merriënboer, J. J. G. (2013). Do Learners Really Know Best? Urban Legends in Education. *Educational Psychologist*, 48(3), 169–183.

Knowles, M. S. (1975). *Self-directed Learning: A Guide for Learners and Teachers*. Cambridge Book Co., New York

Leach, L. (2012). Optional self-assessment: some tensions and dilemmas. *Assess. Evalu. High. Educ*. 37, 137–147.

Leong, H., Yee, C.M., & Kee, C.S. (2019). Flipped Learning to Nurture Self-Directed Learners at Singapore Polytechnic. *Proceedings of the 15th International CDIO Conference*, Aarhus University, Denmark

Lo, C. K. & Hew, K. F. (2017). A critical review of flipped classroom challenges in K-12 education: possible solutions and recommendations for future research. *Research and Practice in Technology Enhanced Learning*, 12(4).

Lo, C. K., Hew, K. F., & Chen, G. (2017). Toward a set of design principles for mathematics flipped classrooms: A synthesis of research in mathematics education. *Educational Research Review*, 22, 50–73.

Malmqvist, J., Edström, K., & Rosén, A. 2020. CDIO Standards 3.0–Updates to the Core CDIO Standards. *Proceedings of the 16th International CDIO Conference*, 60–76.

Mazur, E. (2013). *Assessment: The Silent Killer of Learning* [Lecture]. Derec Bok Center, Harvard University. Retrieved from https://www.youtube.com/watch?v=CBzn9RAJG6Q

Naccarato, E., & Karakok, G. (2015). Expectations and implementations of the flipped classroom model in undergraduate mathematics courses. *International Journal of Mathematical Education in Science and Technology*, 46(7), 968–978.

Nieminen, J. H. (2021), Beyond empowerment: student self-assessment as a form of resistance, *British Journal of Sociology of Education*.

Peters, M. & Prince, M. (2019). Learning and Teaching Engineering Mathematics within an Active Learning Paradigm. *Proceedings of the 15th International CDIO Conference*, Aarhus University, Aarhus

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Ed.), *Handbook of self-regulation: Theory, research, and applications* (452–502). San Diego, CA: Academic.

Rust, C., O'Donovan, B. and Price, M. (2005) 'A social constructivist assessment process model: How the research literature shows us this could be best practice', *Assessment and Evaluation in Higher Education*. Taylor and Francis Ltd , 30(3), pp. 231–240.

Sahlberg, P. (2007). Education policies for raising student learning: The Finnish approach. *Journal of Education Policy*, 22(2), 147–171.

Shepard, L. A. (2005). Linking formative assessment to scaffolding. *Educational Leadership*, 63(3), 70–15.

Swaffield, S. (2011) Getting to the heart of authentic Assessment for Learning, *Assessment in Education: Principles, Policy & Practice*, Vol 18, Issue 4, 433 – 449.

Taras, M. (2009) Summative assessment: the missing link for formative assessment, *Journal of Further and Higher Education*, 33 (1), 57–69.

Toivola, M., Peura, P., & Humaloja, M. (2017). *Flipped learning in Finland*, Edita

Toivola, M. (2016). Flipped learning – Why teachers flip and what are their worries? Experiences of teaching with Mathematics, *Sciences and Technology*, 2(1), pp. 237-250.

Toivola, M. (2020). *Flipped Assessment - A Leap towards Assessment for Learning*, Edita

Toivola, M., Rajala, A. & Kumpulainen, K. (2023). Pedagogical Rationales of Flipped Learning in the Accounts of Experienced Mathematics Teachers, *Pedagogies: An International Journal,* 18:4, 767-787.

Topping, T., Murphy, M., & Saunders, S. (2023). Enhanced Assessment and Learning Through Adaptive Comparative Judgement. *Proceedings of the 19th International CDIO Conference*, NTNU, Norway, 379-389.

Treveyan, J. (2014). *The making of an expert engineer*. New York: CRC Press.

Zimmerman, B. J., and Schunk, D. H. (2011). "Self-regulated learning and performance: an introduction and overview," *in Handbook of Self-Regulation of Learning and Performance*, eds B. J. Zimmerman and D. H. Schunk (New York, NY: Routledge), 1–14.

BIOGRAPHICAL INFORMATION

Marika Toivola works as a Senior Lecturer at the Satakunta University of Applied Sciences, Pori, Finland and is a Doctoral Researcher at University of Helsinki, Finland. She is a well-known pioneer in Flipped Learning and Flipped Assessment and won the Mathematics Teacher of the Year 2019 in Finland Award. Her passion is in research-based teaching development. Toivola has published several professional books for teachers and statements on self-regulated learning and assessment for learning.

Corresponding author

Marika Toivola
Satakunta University of Applied Sciences
Satakunnankatu 23
28130 Pori, FINLAND
Marika.Toivola@samk.fi