CHALMERS

Proceedings

Chalmers Conference on Teaching and Learning 2025

KUL 2025

22 APRIL 2025 JOHANNEBERG CHALMERS
KONFERENS OM UNDERVISNING OCH LÄRANDE (KUL2025)

PROCEEDINGS CHALMERS
CONFERENCE ON TEACHING AND LEARNING 2025

Chalmers University of Technology, 22 April 2025
Organised by Chalmers University of Technology, Sweden

Venue: Chalmers Conference Center, Johanneberg, Gothenburg, Sweden

ISBN: 978-91-88041-64-7
DOI: 10.5281/zenodo.20134179

Editors: Magnus Gustafsson, Per Lundgren, Lena Peterson och Christian Stöhr
Graphic design: Dahlbäck/Söderberg
Photo: Henrik Sandsjö

Table of Content

About Chalmers Conference on Teaching and Learning (KUL) 5
Committees, groups, and roles 6
Keynote speaker 7

Papers

Ensuring student completion of preparatory work in flipped classes – Example from a TRACKS course
Christophe Demazière 9

Improving Coding Assignments with Partial Auto-Grading and Immediate Feedback: Report from an intervention
Marco L. Della Vedova 19

Participation in close-to-practice educational research – researchers' perspective
Samuel Bengmark, Philip Gerlee and Jonathan Weidow 29

Work in progress

Swinging Pendulums on the Cloud: Digitization of Simulation & Experimental Infrastructure for Feedback-based Active Learning
Shivesh Kumar 37

Workshops

Assessing student-centred learning without a written exam? A discussion on student-centred learning and continuous examination
Ana Carolina Bertassini and Anna Nyström Claesson 49

About Chalmers Conference on Teaching and Learning

KUL, Chalmers Conference on Teaching and Learning, is held yearly and started in 2011. The aim is to contribute to the quality of the entire educational activity at Chalmers, i.e., undergraduate education, postgraduate education, and collaboration, by promoting collegial conversation about learning and teaching.

KUL offers an opportunity for the exchange of experience regarding program and course development and is an arena for pedagogical development and qualification.

The idea is that the Keynote and other sessions should contribute to the collegial conversation and inspire further exchange, also between the annual conferences.

These proceedings gather the works of Chalmers teachers that typically emerge directly from the classroom. They provide insight into the lived experiences of teachers who actively seek to refine their practices and reflect on their pedagogical challenges. Hence, the contributions offer examples of the scholarship or teaching and learning (SoTL) and suggest avenues for subsequent educational development research.

The contributions presented here are aimed to serve as valuable starting points for dialogue, inspiration, and ongoing collegial learning.

Summary in brief for KUL2025

Date: 22 April 2025
Venue: Chalmers Conference Center, Johanneberg, Gothenburg, Sweden

Number of participants: 138
Number of conference contributions: 34
Number of contributing authors: 52
Number of plenary sessions: 1
Number of parallel sessions: 20, over five parallel sessions
Average rating: 4.2 out of 5 (31% response rate)

Conference committee and evaluators of contributions

Conference committee and review group for conference contributions

Lena Peterson, MC2, (chair)
Per Bjerkeli, Head of Program, Global Systems, SEE
Malin Blomqvist, Management and university joint support
Karl de Fine Licht, educational developer, TME
Magnus Gustafsson, expert, CLS
Beata Holmlid, representative of the Chalmers student union
Verena Siewers, vice-prefect for undergraduate education, LIFE
Christian Stöhr, expert, CLS

Review group for proceeding contributions

Magnus Gustafsson
Christian Stöhr

Session chairs

Jimmy Ehnberg
Laura Fainsilber
Per Lundgren
Lena Peterson
Verena Siewers
Christian Stöhr
Marta Vetier

Conference administration

Anna Bergius
(Internal Communications), conference administrator and web manager

Keynote speaker 2025
The case for contextualized and interdisciplinary engineering education: From promises to practices

Employers, educators, and the public need engineers who respond knowledgeably and ethically to the needs of industry, stakeholders, and the communities their work affects immediately and in the future, yet engineering courses often abstract technical knowledge from the context-laden, real-world engineering problems that require definition and sound solutions. Moreover, experience and research suggest that the interdisciplinary and contextual competences needed for responsive and responsible engineering practice take time to develop and deepen.

At Chalmers, Tracks options and thoughtfully constructed design and project courses introduce students to interdisciplinary collaboration and problem-solving experiences. Extra-curricular activities can further engage students, although engagement is optional. A question thus arises: How might engineering programmes that seek to cultivate interdisciplinary and habits of mind and contextualized problem solving further integrate this learning into the formal curriculum and promote understanding of the sociotechnical dimension of engineering work? In this talk I draw on research on engineering education to make a case for contextualized and interdisciplinary engineering learning experiences. I also offer suggestions for educational practices and policies that support an educational programme that quite intentionally prepares engineers to work in ways that are responsive to the needs of employers and stakeholders and committed to contributing to an engineered world in which individuals and communities can live and thrive.

Lisa R. Lattuca
Professor at the Center for the Study of Higher and Postsecondary Education at the University of Michigan.

Papers

Ensuring student completion of preparatory work in flipped classes – Example from a TRACKS course

Christophe Demazière
demaz@chalmers.se

August 23, 2025

Abstract

Flipped classroom pedagogy is based on students completing preparatory self-studies before participating in more active forms of learning. Unfortunately, students do not always adhere to the necessity to have completed the preparatory work. We present hereafter some course design principles ensuring that students come to class prepared. Those principles were applied to a Master level course offered at Chalmers University of Technology, as part of the TRACKS initiative. The effectiveness of those principles is analysed separately via completion rates of the preparatory activities, success rates of those, and overall course performance. It is demonstrated that the design principles resulted in a high student engagement throughout the entire course, both during the preparation for the more active forms of learning and during such sessions. As a result, students obtained high grades, and the overall course satisfaction was also high.

Sammanfattning

Flipped classroom-pedagogik bygger på att den studerande genomför förberedande självstudier innan de deltar i mer aktiva former av lärande. Tyvärr håller studenterna inte alltid fast vid nödvändigheten av att ha slutfört förarbetet. Vi presenterar några principer för kursdesign som säkerställer att studenterna kommer förberedda till lektionen. Dessa principer tillämpades på en kurs på masternivå som gavs vid Chalmers tekniska högskola, som en del av TRACKS-initiativet. Effektiviteten av dessa principer analyseras separat genom slutförandegraden av de förberedande aktiviteterna, framgångsfrekvenser för dessa och övergripande kursprestationer. Det visas att designprinciperna resulterade i ett högt studentengagemang genom hela kursen, både under förberedelsen för de mer aktiva lärandeformerna och under sådana sessioner. Som ett resultat fick studenterna höga betyg, och den övergripande kursnöjdheten var också hög.

Keywords: *flipped classroom; asynchronous vs. synchronous activities; active learning.*

1. Introduction

Flipped classrooms have received some increased attention in the past 15 years in the area of engineering education (Bishop & Verleger, 2013; O'Flaherty & Phillips, 2015). Flipped classroom pedagogy consists in dividing student learning sequence in a set of asynchronous activities and synchronous activities, respectively. The asynchronous activities are self-paced activities, typically arranged in forms of reading, watching some videos/video lectures, possibly complemented by quizzes and other activities. Those activities are organized ahead of and as a preparation to the synchronous activities. Although teachers monitor those activities and students can contact the teachers for getting help or in case of questions, there is no direct "face-to-face" interaction between the teachers and students during this phase. The synchronous activities, on the other hand, rely on more involved activities requiring the simultaneous interactions between the teachers and the students. In most flipped set-ups, the synchronous activities are organized onsite. They can also be arranged fully online or in a mixed manner, i.e., having some students on campus and some students online (Felder & Rebecca, 2024; Stöhr, Demazière & Adawi, 2020; Demazière et al., 2024).

The asynchronous phase can be seen as a preparatory phase, during which students acquire the necessary knowledge to later work on more advanced activities during the synchronous phase. Whereas the asynchronous phase requires very little teacher support and can thus be offered as self-paced activities, teacher presence and interactions with the students is essential for the synchronous phase. The activities during this second phase are based on student themselves undertaking such activities, typically in groups. Those activities should be developed using active learning techniques (Freeman, et al., 2014). Flipped classrooms can be seen as making more time available in the course curriculum for such sessions, by removing from the classroom activities focusing on acquiring the necessary knowledge (such as via traditional lectures). As a result, teachers can focus the classroom time on activities during which students apply their knowledge and develop their skills, possibly complemented by interactive sessions with focused lectures.

Flipped classes have been demonstrated to lead to higher student engagement, deeper approach to learning, better fulfilment of the learning objectives and better student performance (Karabulut-Ilgu, Cherrez & Jahren, 2018). Nevertheless, for flipped classroom pedagogy to be beneficial to students, students must complete the necessary preparatory asynchronous work ahead of the synchronous activities. Lack of student compliance with the necessity to carry out the preparatory work is mentioned as one of the main difficulties encountered by teachers when implementing flipped classrooms (Pfennig, 2024).

Failing to prepare for the active learning sessions results in uneven student readiness to undertake the activities planned by the teachers during the sessions, partially or fully jeopardising those activities. This often leads to teacher frustration, as the teacher needs to on-the-fly change the planning of the synchronous activities. In some cases, this frustration also results in teachers giving up flipped classroom pedagogy and reverting to more traditional forms of teaching.

Pfennig (2024) presents some mitigation measures, some of them can be implemented during the synchronous sessions in case some students did not complete the necessary preparatory work. Building upon some of these measures, we additionally focus, in the work reported hereafter, on principles, from a course design perspective, ensuring that students come to class prepared. In (Demazière et al., 2024), some course design features are presented, in which the necessity to complete the asynchronous work is an eligibility condition to participate to the synchronous activities. Nevertheless, this set-up is applied to courses in which the asynchronous phase is entirely carried out before the synchronous activities, concentrated on 5-10 consecutive days, take place. For a course requiring more intertwined sequences of asynchronous and synchronous activities, a different approach is necessary.

In this study, the efficacy of the course design principles implemented to ensure that students come to the synchronous sessions prepared is investigated. More specifically, the work reported hereafter addresses the three following questions in the adopted course set-up:

- Do the students complete the self-paced online preparatory work each week, to which amount, and on time?

- How do the students perform on asynchronous quizzes used in the asynchronous sessions?
- How do the students perform on the course overall?

The course considered hereafter was a new course, and there is thus no possible comparison with earlier iterations of the course.

The paper is structured as follows. The course set-up is first presented. The generic features of the course with respect to flipped classrooms are introduced, before looking at the specificities of the course design implemented in the present course. Thereafter, the methodology to address the above research questions is discussed, followed by the results of the analysis. The paper ends with some discussion and conclusions.

2. Course set-up and features

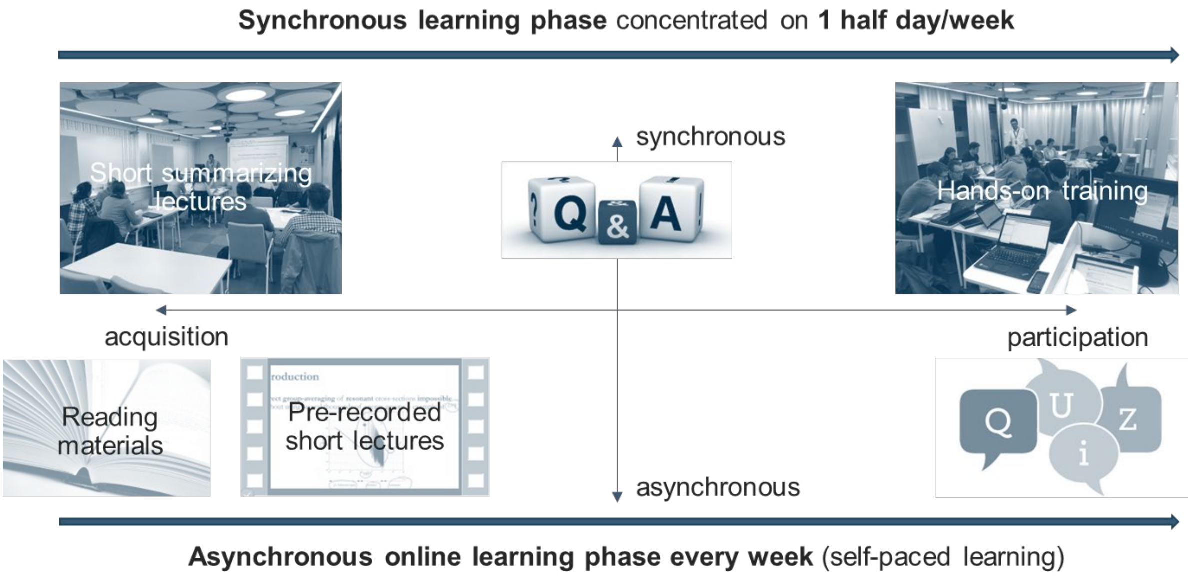
The course was an elective course in nuclear engineering and offered at the master level. Beyond MSc students, the course was also open to PhD students, professionals and ERASMUS students. For the professionals, the course could be followed entirely online. The course iteration reported hereafter had 45 active students and was given during the academic year 2024/2025. Four of those students were professionals and followed the course entirely online. Out of the remaining 41 students, seven were PhD students, 33 were MSc students and one was a BSc student. The overall pedagogical concept for the course is based on a flipped classroom pedagogy, which consists in an asynchronous phase followed by a synchronous phase. Because of the remote location of some of the course attendees, the synchronous sessions were simultaneously offered to the onsite students and live-streamed to the remote students (hybrid set-up). The synchronous sessions, although not mandatory, were also recorded and made available to all students short after each session. This made it easier for students missing a session to catch up. The learning sequence of asynchronous/synchronous activities was repeated every week over a period of 14 study weeks (spreading over nearly four months).

The asynchronous phase consisted of (a) reading a set of handbooks specifically written for the course, (b) watching some short summarizing video lectures aimed at capturing the main concepts presented in the handbooks, and (c) answering some quizzes associated to each of the video lectures.

The synchronous phase consisted of (a) short summarizing lectures, (b) the possibility for students to ask question and clarifications, and (c) more advanced hands-on activities that the students needed to work on in groups of four to five students. Various activities were undertaken over the entire duration of the course: problem-based solving, demonstrations, laboratory exercises, and workshops.

The entire learning sequence the students were exposed to each week is presented in Fig. 1.

Figure 1: Overview of the various course elements offered in the course, categorized along two dimensions following Hrastinski (2008): acquisition versus participation, and asynchronous versus synchronous. The asynchronous learning phase (bottom row of the figure) had to be completed before the synchronous learning phase (top row of the figure) each week of the course.



The additional course designs principles that were implemented are summarized as follows:

- Instructions about the course design principles and how the students should work were distributed to the students about one month before the start of the course.

- Flipped classroom pedagogy was introduced in the instructions. Furthermore, the first introductory synchronous session was partly devoted to explaining flipped classrooms and how students should work. Using research-based evidence, for instance from Felder and Rebecca (2024), the advantages for students were detailed (focusing on coming to the synchronous sessions much better prepared, better managing own study time, and learning better and more efficiently).

- All teaching resources during both the asynchronous and synchronous phases were delivered via the Learning Management System (LMS) used for the course. A highly structured course page was designed, where each week a learning sequence for the asynchronous work had to be followed. The asynchronous work typically consisted of reading some handbook written for the course, watching some short summarizing videos and answering some corresponding quizzes. A locking mechanism forced the students to study those resources in a predefined order.

- Continuous assessment was used throughout the course and there was no final examination. The asynchronous activities had a weight of 25% on the final course grade, with correspondingly a weight of 75% on the activities undertaken during the synchronous sessions. The activities in those sessions typically consisted of group work and discussions focusing on small projects, quizzes, modelling/computational tasks, laboratory exercises and workshops.

- Overall course grades were updated on-the-fly as students completed both asynchronous and synchronous activities.

- A variety of types of quizzes was used during the asynchronous phase: true/false, multiple choice/answers, categorisation, matching, ordering, hot spot. Automatic feedback was provided to the quizzes, both for correct and incorrect answers. The quizzes were graded. The quizzes had a limited number of attempts, and the quiz scores reflected the number of attempts. For some more demanding quizzes, an unlimited number of attempts was permitted and the highest score retained. The deadlines on the quizzes were at the end of the day before the start of the associated synchronous sessions. Missing the deadline resulted in the quizzes no longer being available to students and thus in lower scores on the asynchronous activities.

- The completion rates on the quizzes and student success rates were monitored by the teacher. Students who did not complete the quizzes on time were contacted by the teacher. Also, depending on how students performed on the quizzes, additional explanations were provided during the following synchronous session, following the principles of Just-in-Time-Teaching – JiTT (Novak et al., 1999). Students were encouraged to put questions during the asynchronous phase.

- Student scores on both the asynchronous and synchronous activities during the first month of the course were used to nominate 10 students who thereafter could participate in laboratory exercises at a research facility abroad, the other 35 students participating remotely instead.

- After each synchronous session, a message was sent by the teacher summarizing what needed to be done prior to the next synchronous session.

3. Methodology

As all teaching resources considered in the course were provided to the students via the LMS, the corresponding activities can be monitored. As earlier mentioned, the asynchronous work typically consisted of reading sets of handbooks, watching videos and answering quizzes. Although the use of the handbooks and videos is monitored by the LMS, only access to those resources is recorded. Accessing a handbook or a video does not guarantee that students do read the handbook or watch the video. Quizzes are considered as a more reliable source of student activity, also considering that the quizzes had been thoroughly developed and tested before being offered to students. Not only the occasions at which activities on the quizzes are started are logged by the LMS, the LMS also keeps track of when the quizzes are submitted. Furthermore, whether students obtained the correct answers or not on the quizzes is recorded.

Consequently, analytics on quizzes are used hereafter to quantity the effect of the course design principles on the asynchronous preparatory work. For each student enrolled in the course and for each week of the course, the logged times at which the quizzes were completed were extracted. Likewise, the corresponding grades were recorded. For true/false quizzes, only one attempt was allowed, whereas for all other types of quizzes, multiple attempts were allowed and an average grade reflecting the number of attempts was applied. In case a quiz was not submitted, the corresponding score was set to zero.

None of the above data is readily available from the LMS. Significant post-processing of the logged student activities was required at the student level to estimate the above metrics.

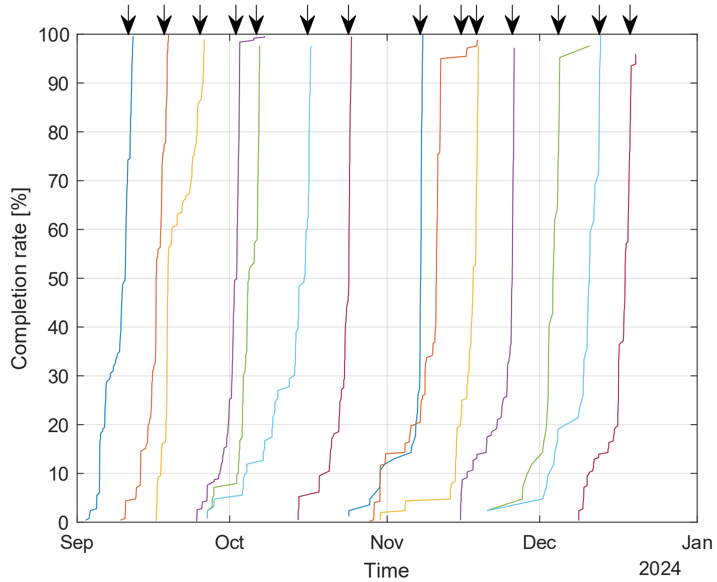
The analysis of weekly completion rate on the quizzes allows addressing the first research question. The analysis of weekly success rate on the quizzes can be used for the second research question. Based on student individual completion and grades on the quizzes, a weekly timeline of cumulative completion rates and grades, respectively, was computed. For each week, the cumulative quantities were estimated as follows. If all students had completed all quizzes for a given week, the cumulative completion rate for that week would be 100%. Likewise, if all students had correctly answered all quizzes in only one attempt for a given week, the score for that week would be 100%. The overall course scores readily available from the LMS were utilized to answer the third question.

4. Results

The course ran on a period of 14 weeks, with a varying number of quizzes per week. Across those weeks, the cumulative completion rates for all students collectively vary from 90.7% to 100%, whereas the cumulative scores vary from 73.1% to 97.2%. Those measures were estimated by considering the completion rates and scores, respectively, for every week but for all students together. 100% would therefore mean that all students either completed all quizzes for a given week or obtained the maximum score on the quizzes for that week. Variations between students exist. Among the 45 active students, three students gave up and did not complete any course activity after roughly the mid-course. Retaining the 42 active students, the metrics improve: the cumulative completion rates for all students vary from 95.9% to 100% across all weeks, whereas the cumulative scores vary from 77.4% to 98.4%.

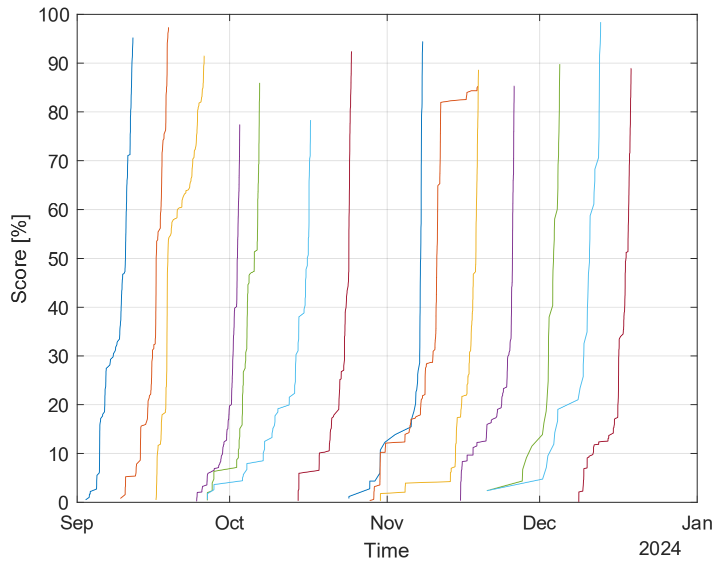
The timeline of the completion rate and of the scores on the asynchronous quizzes are reported in Figs. 2 and 3, respectively, throughout the entire duration of the course (from beginning of September to end of December 2024). On those figures, each curve represents for a given week the evolution of the completion rate on all quizzes (Fig. 2) or the corresponding score (Fig. 3) pertaining to that week for all students together. 100% would therefore mean that all students either completed all quizzes for a given week or obtained the maximum score on the quizzes for that week. One can also verify on Fig. 2 that the cumulative completion rate at the end of each week varies from 95.9% to 100% across all weeks, as earlier stated. Likewise, one can also observe on Fig. 3 that the cumulative scores at the end of each week varies from 77.4% to 98.4% across all weeks.

Figure 2: Timeline (from beginning of September to end of December 2024) of the evolution of the completion rate on the asynchronous quizzes for the active students. The arrows represent the deadlines at which the quizzes were due.



As Fig. 2 demonstrates, the asynchronous work was carried out per the agreed deadlines for most students, as the cumulative completion rates reached nearly 100% by the agreed deadline every week. By design, the quizzes became unavailable after the deadline. In most cases, students completed the quizzes on time. In a few rare occasions, deadline extensions were agreed upon on a case-by-case basis with some students and for personal reasons.

Figure 3: Timeline (from beginning of September to end of December 2024) of the evolution of the score on the asynchronous quizzes for the active students.

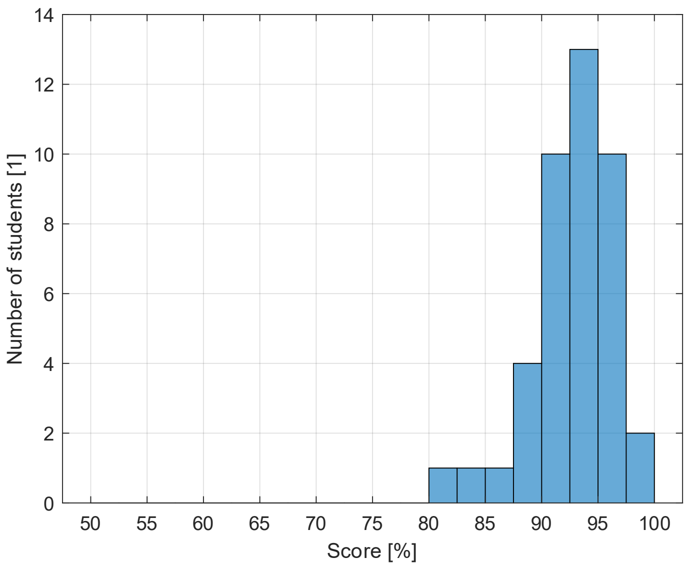


Also, it can be noticed from Fig. 3 that a high success rate on the asynchronous quizzes was obtained, with a few exceptions, as the cumulative scores reached roughly 80 to nearly 100% every week. The generally high success rate demonstrate that the students properly assimilated the concepts they were exposed to via the handbooks and videos. In the cases when the success rates were lower, the teacher returned to those quizzes and clarified the reasoning behind the quiz questions. Following the principles of JiTT, the student answers to the asynchronous quizzes were studied by the teacher ahead of each synchronous session. This allowed the teacher to fine tune the sessions on the student needs and to identify students possibly in difficulty.

Concerning the synchronous sessions, the cumulative success rates on each part of the course varied from 85.0% to 100%. Because of the high success rates on both the asynchronous and synchronous activities and their alignment with the intended learning outcomes of the course, it can be claimed that the students achieved those objectives to a very large extent.

Concerning the overall course scores, they vary between 81.1% and 98.6% of the maximum course score. The distribution of the student scores is represented in Fig. 4. Considering that the synchronous activities represent 75% of the overall course score, it can be concluded that students correctly assimilated the course concepts during the asynchronous phase. This allowed them to perform extremely well on the more challenging synchronous activities, as the above cumulative success rates on those activities demonstrate. It is also worth noting that both the onsite and online students performed equally well on the course (data not reported here for the sake of brevity).

Figure 4: Distribution of the overall course scores.



5. Discussion and conclusions

As the above measures demonstrate, the course design features implemented for ensuring that students come to the synchronous sessions prepared worked, with a high completion rate on the asynchronous quizzes, from the early start of the course till the end. The fact that the scores during the first month of the course were used as a basis to determine which students would be able to participate onsite to the laboratory exercises abroad was definitely an important ingredient for reaching such a high student engagement in the course. Other motivating elements could be considered in other course settings, where extra student activities/engagement could be for instance rewarded. The incentive in studying the course elements early also resulted in the students adopting flipped classroom pedagogy from the beginning of the course and appreciating the benefits of the course set-up. This led to the students continuing to work throughout the entire course with the same level of dedication. As correctly answering the quizzes requires studying some additional materials (reading the handbook and/or watching the videos), the high success rates on the quizzes imply that the students indeed completed the preparatory work. This resulted in students coming to the synchronous interactive sessions well prepared. Beyond completion of the quizzes, the students thus successfully learned the basic concepts of the course, as the scores on the quizzes and overall course score demonstrate. The students worked on the course and the quizzes contentiously and thoroughly during the asynchronous phase.

Since the course is elective, students were expected to be highly motivated for the course. Nevertheless, high student motivation for the course may only partly explain the large adoption to the course design principles. Students may have chosen to only pass the course, especially considering that they immediately knew the evolution of their grades while completing activities. As the course results demonstrate, students obtained extremely high scores on the course, thus largely exceeding just the passing grade.

It should be noted that grading the quizzes and using them as part of the final course grade has the drawback of not letting the students train during the asynchronous phase. As the quizzes are graded and thus used as continuous assessment, students are very cautious before answering the quizzes. Although resulting in a higher seriousness in answering the quizzes, the students do not have the possibility to use the quizzes as formative assessments. This could be addressed by adding another series of quizzes or example problems on which the students could train.

Some of the course design principles presented in this work could be used as a source of inspiration for other teachers planning to adopt flipped classrooms or struggling with students to complete the necessary preparatory work on time in flipped courses.

The continuous feedback the students receive on all learning activities, beyond their formative nature, allow the students to see their progression towards passing the course, adding an extra ingredient for motivating them to complete the tasks. Finally, through careful course planning and design, the active learning activities proposed during the synchronous sessions resulted in students able to solve advanced problems.

Some aspects of the course design are planned to be modified in the subsequent editions of the course. First, more time will be given to the students to brainstorm and work in groups before support is provided by the teachers. In this respect, more information will be provided ahead of the sessions, so that students get a better idea of the problems and activities proposed during the sessions. Also, this will allow them to better prepare during the asynchronous phase for those activities. A more in-depth analysis of possible mismatch in results between the asynchronous quizzes and the synchronous activities would also allow assessing the efficacy of the quizzes, possibly leading to some re-design of those. Finally, some students may have benefited from their peers when working in groups. More individual forms of assessment of personal contributions to group work will be considered.

References

Bishop, J., & Verleger, M. A. (2013). The Flipped Classroom: A Survey of the research. 23.1200.1-23.1200.18. https://peer.asee.org/the-flipped-classroom-a-survey-of-the-research

Demazière C., Stöhr C., Zhang Y., Cabellos O., Dulla S., Garcia-Herranz N., Miró R., Macian R., Szieberth M., Lange C., Hursin M. & Strola S. (2024). Enhancing higher education through hybrid and flipped learning: Experiences from the GRE@T-PIONEeR project. Nuclear Engineering and Design, 417, 113028. https://doi.org/10.1016/j.nucengdes.2024.113028

Felder R. M., & Rebecca B. (2024). Teaching and Learning STEM : A Practical Guide, John Wiley & Sons, Incorporated

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences, 111(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Hrastinski, S. (2008). Asynchronous and synchronous e-learning. Educause Quarterly, 31 (4), 51–55. https://er.educause.edu/articles/2008/11/asynchronous-and-synchronous-elearning

Karabulut-Ilgu, A., Cherrez, N. J., & Jahren, C. T. (2018). A systematic review of research on the flipped learning method in engineering education. British Journal of Educational Technology, 49(3), 398–411. https://doi.org/10.1111/bjet.12548

Novak, G. M., Patterson E., Gavrin A., & Christian W. (1999). Just-in-Time Teaching: Blending Active learning with web technology. Upper Saddle River, NJ, Prentice Hall.

O'Flaherty, J., & Phillips, C. (2015). The use of flipped classrooms in higher education: A scoping review. The Internet and Higher Education, 25, 85–95. https://doi.org/10.1016/j.iheduc.2015.02.002

Pfennig, A. (2024). The challenge of unprepared students in inverted classroom teaching scenarios. In: 10th International Conference on Higher Education Advances (HEAd'24). Valencia, 18-21, June 2024. https://doi.org/10.4995/HEAd24.2024.17067

Stöhr C., Demazière C., & Adawi T., (2020). The polarizing effect of the online flipped classroom. Computers and Education, 147, 103789. https://doi.org/10.1016/j.compedu.2019.103789

Improving Coding Assignments with Partial Auto-Grading and Immediate Feedback: Report from an intervention

Marco L. Della Vedova
marco.dellavedova@chalmers.se

September 8, 2025

Abstract

This paper presents a case report on the use of partial auto-grading with immediate feedback in the course *Introduction to Artificial Intelligence* (MMS131) at Chalmers University of Technology. The intervention aimed to reduce grading workload while providing students with timely, formative feedback on coding assignments. Evaluation was carried out by comparing the 2023 course iteration (manual grading) with the 2024 iteration (auto-grading integrated into Jupyter notebooks). The case showed reduced grading time, shorter feedback delays, and lower resubmission rates, alongside generally positive perceptions from both students and teaching staff. Challenges included installation issues and occasional concerns about the clarity of automated tests. We discuss the pedagogical and operational implications of these findings and reflect on design choices such as scaffolding with a preliminary assignment and balancing visible and hidden tests.

Abstract

Denna artikel presenterar en fallstudie om användningen av partiell auto-rättning med omedelbar återkoppling i kursen *Introduction to Artificial Intelligence* (MMS131) vid Chalmers tekniska högskola. Interventionen syftade till att minska arbetsbördan för rättning samtidigt som studenterna fick snabb, formativ återkoppling på sina programmeringsuppgifter. Utvärderingen genomfördes genom att jämföra kursupplagan 2023 (manuell rättning) med upplagan 2024 (auto-rättning integrerad i Jupyter notebooks). Fallstudien visade på minskad rättningstid, kortare väntetider för återkoppling och färre om-inlämningar, tillsammans med överlag positiva upplevelser från både studenter och lärare. Utmaningar inkluderade installationsproblem och viss oro kring tydligheten i de automatiska testerna. Vi diskuterar de pedagogiska och praktiska implikationerna av resultaten och reflekterar kring designval såsom användning av en inledande uppgift för att underlätta introduktion samt balansen mellan synliga och dolda tester.

Keywords: *coding; assignments; auto-grading; immediate feedback*

1 Introduction

In the course Introduction to Artificial Intelligence (MMS131) at Chalmers University of Technology, students are required to complete assignments that combine programming tasks with written, pencil-and-paper exercises. The grading process for these assignments was particularly critical and labor-intensive due to two main factors: (i) The course consistently enrolls a large number of students, i.e., about 200; (ii) All assignments were manually graded by teaching assistants (TAs) and lecturers.

This traditional grading approach presents challenges that affect both teaching and learning:

- **Delayed feedback**: Manual grading of all exercises for all students leads to significant delays in returning grades and feedback.
- **High workload**: The grading burden on teaching staff is substantial, reducing the time available for other pedagogical activities.

To address these challenges, this work explored introducing an auto-grading system for the programming exercises. This system is designed to provide immediate, automated feedback to students, thereby enabling them to identify and correct errors in a timely manner. In parallel, it aims to reduce the grading workload for teaching staff, without compromising the quality or integrity of assessment. Importantly, the intervention does not seek to fully automate the grading process. Instead, auto-grading is used as a pedagogical aid to support student learning and foster iterative problem-solving. Final assessment decisions remain the responsibility of the teaching staff, ensuring alignment with learning outcomes and maintaining academic standards.

This paper reports on the design, implementation, and evaluation of this intervention, addressing the following research questions:

RQ1: What changes in grading workload and feedback turnaround are observed after introducing partial auto-grading with immediate feedback, compared with the previous course iteration?

RQ2: What changes occur in resubmission rates and in assignment marks on the redesigned tasks?

RQ3: How do students and TAs perceive the usefulness and usability of the auto-grading workflow (e.g., confidence, error detection, clarity of tests, installation/setup)?

It is worth noting that this course is a mandatory third-year course for students in the Mechanical Engineering bachelor's program and an elective for those in Industrial Design Engineering. Additionally, it is a popular choice among exchange students, with approximately 20–30 enrolling each year. Students from these backgrounds often do not identify as "programmers," and, according to the initial course survey, around half express limited interest in coding. This underscores the importance of designing structured assignments that guide students step-by-step through the problem-solving process.

2 Related work

There is a quite extensive literature about auto-grading, or *Automatic Assessment Systems*. It is established among computer science teachers that "manually grading programming assignments is time consuming and tedious, especially if they are incorrect and incomplete" (Nikhila & Chakrabarti, 2022, Abstract), negatively affecting the teaching experience. Automatic assessment systems have been developed to alleviate these issues, improving both the efficiency and effectiveness of grading.

Early studies, such as (Douce, Livingstone, & Orwell, 2005; Edwards & Pérez-Quiñones, 2008), show that automatic grading systems are important because they provide quick feedback to students. This immediate feedback helps students find and fix mistakes quickly, which is important for learning. A more recent study recognizes the value of auto-grading, but alerts that "students are not likely to develop testing skills since they over-rely on the feedback from the auto-grader." (Mitra, 2023, Abstract).

Recent improvements have focused on making auto-grading systems better for teaching (Aldriye, Alkhalaf, & Alkhalaf, 2019). In particular, Hegarty-Kelly and Mooney (2021, Abstract) pointed out that "lecturers can also provide instant feedback to students while keeping track of their progress and identifying where the gaps in students' knowledge are." This ability to monitor and adapt to student performance in real-time is a major benefit of auto-grading systems.

Paiva, Leal, and Figueira (2022, Abstract) discussed how assessing a program is more complex than just checking if it works correctly. They mentioned that "assessing a program is considerably more complex than asserting its functional correctness, as the proliferation of tools and techniques in the literature over the past decades indicates. Program efficiency, behavior, and readability, among many other features, assessed either statically or dynamically, are now also relevant for automatic evaluation." This shows that relying entirely on automated systems for grading can lead to misjudgment.

In summary, research on auto-grading for programming assignments covers many areas and shows ongoing improvements in technology and teaching methods. The main advantages are the immediate feedback for students and the speedup in grading for teachers. The main disadvantages are that students tend to over-rely on the feedback from the auto-grader, and it can be difficult for teachers to select the appropriate tool given the proliferation of these systems.

3 Methodology

This study evaluates the impact of integrating auto-grading into coding assignments through a comparative analysis of two consecutive iterations of the Introduction to Artificial Intelligence course. The 2023 iteration served as the baseline, utilizing traditional manually graded assignments, while the 2024 iteration featured redesigned assignments incorporating auto-grading with immediate feedback. To assess the impact of the intervention, we report quantitative data and feedback from students and TAs. This provided insight into both measurable outcomes and user experiences.

The methodology is organized into two subsections: Section 3.1 describes the implementation of the intervention, while Section 3.2 outlines the evaluation approach.

3.1 Implementation

The course includes four assignments in total, but only the first two were redesigned to incorporate auto-grading. The remaining two assignments are managed by other instructors and were therefore not modified as part of this intervention. As a result, students experienced both the traditional and the redesigned assignment formats within the same course, providing a natural point of comparison.

Tool Selection Process To select a suitable auto-grading solution, a comparative review of existing tools was undertaken. Factors considered included compatibility with Jupyter notebooks, ease of integration with the existing teaching infrastructure, i.e., Canvas, support for partial grading and detailed feedback, and the overall user experience for both instructors and students. As a result of the review, **Otter-Grader** (Pyles & UC Berkeley Data Science Education Program, 2024) has been selected as the auto-grading tool. The main reason is that, differently from cloud-based solutions, it does not require students to register to a third-party service. Moreover, Otter-Grader aligns with the philosophy of the new assignment design, that is to guide students towards the solution of the problems with intermediate steps. However, the tool does have some drawbacks. Students are required to install additional software—albeit in the form of a standard Python package—and the system can be somewhat complex for instructors to configure and use effectively.

Design of new assignments Following the selection of the auto-grading tool, the assignments were redesigned to incorporate automated feedback using Jupyter notebooks integrated with Otter-Grader. The notebook is structured into distinct questions that can build on one another, and tests are implemented for selected questions. These tests are designed to evaluate the functionality of the students' code, similarly to *unit testing* in software engineering. Some of the tests are visible to students, providing immediate feedback during their development process. Others remain hidden and are used exclusively by the teaching staff for verification purposes[1]. Some notebooks were deliberately scaffolded with step-by-step instructions and partially completed code templates, guiding students towards the intended solution. Others were presented in a more open format, leaving students with greater freedom in how to design and implement their solutions.

The implementation effort involved the development of two new assignments, each consisting of three problems, resulting in a total of six Jupyter notebooks built around the new structure. Several problems were adapted from previous years with minor modifications, while others were newly designed. In addition, a new assignment, called *Assignment n.0*, was introduced to familiarize the students with the system. As an example, Figure 1 shows the Jupyter notebook used for Assignment 0.

The overall redesign effort was somewhat greater than the usual yearly updates to the assignments, since additional attention had to be given to designing and validating the automated tests, but the extra workload was moderate and manageable.

Instructions to students Clear instructions were given to the students on how to install Otter-Grader and how to use it. For this purpose a "Getting started" page was published on Canvas, where (among other things) it was stated:

"For assignments 1 and 2 we will experiment the use of a system to give you immediate feedback on your solutions. Note that the automatic feedback is not comprehensive and the final grading will be done by the teachers. So, if you get all tests to pass, it does not mean that you have a perfect solution. The automatic feedback is only meant to help you testing your code and to give you some guidance on how to solve the problems."

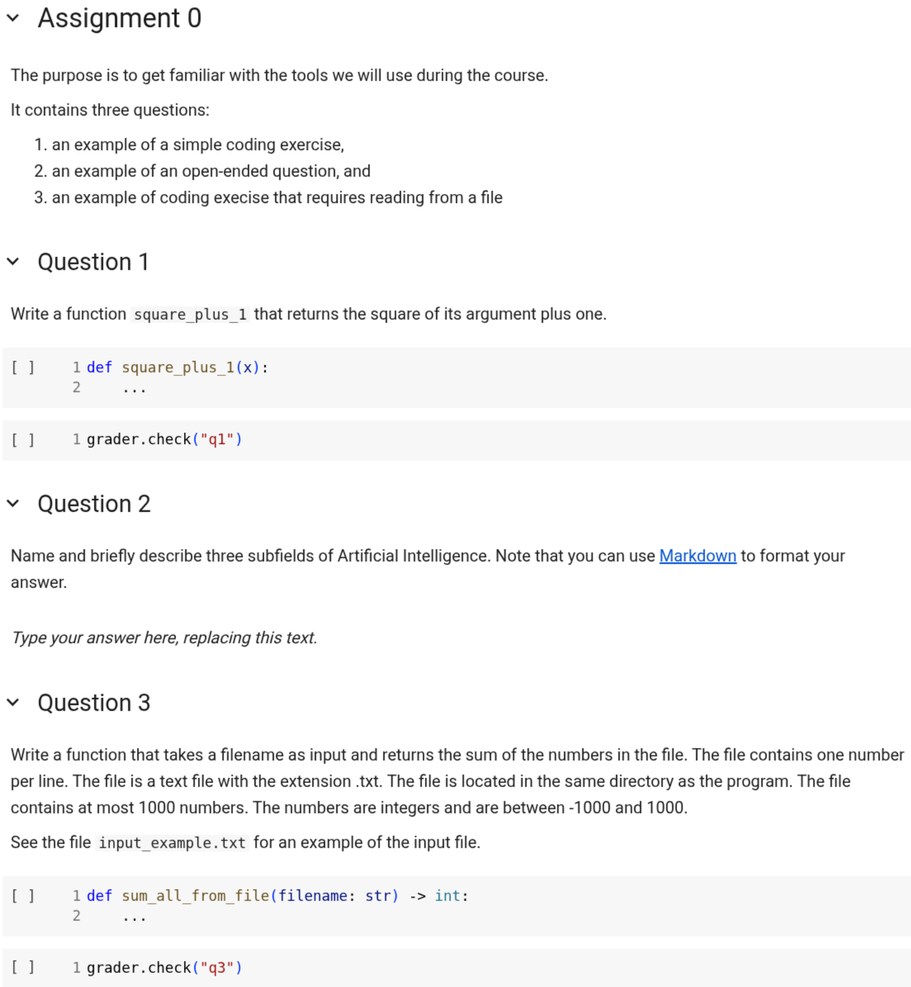
Workflow Each assignment followed a structured workflow. Students first received a Jupyter notebook containing the assignment hand-out. While working on the tasks, they could execute their code locally and obtain immediate feedback from the automated checks, allowing them to verify—before submission—that their solution passed all available tests. Once satisfied, students submitted their notebooks through the course platform (Canvas). The teaching assistants then reviewed the submissions: they re-ran the notebooks with the autograder (with additional tests), but also performed a manual inspection to evaluate aspects not fully captured by automation, such as clarity, readability, and efficiency of the code. Grades and written feedback were then provided, and students with insufficient or incorrect solutions were asked to resubmit. In this way, the auto-grader supported formative learning while final evaluation remained under instructor control.

3.2 Evaluation of the intervention

The evaluation combined course metrics with feedback from students and teaching assistants.

[1]See `https://otter-grader.readthedocs.io/en/latest/test_files/python_format.html` for more information about how to write tests in Otter-Grader.

Figure 1: The Jupyter notebook for Assignment 0. Question 1 and 3 provide automated feedback with `grader.check`; while question 2 does not.



Assignment 0

The purpose is to get familiar with the tools we will use during the course.

It contains three questions:

1. an example of a simple coding exercise,
2. an example of an open-ended question, and
3. an example of coding execise that requires reading from a file

Question 1

Write a function `square_plus_1` that returns the square of its argument plus one.

```
[ ]  def square_plus_1(x):
         ...
```

```
[ ]  grader.check("q1")
```

Question 2

Name and briefly describe three subfields of Artificial Intelligence. Note that you can use Markdown to format your answer.

Type your answer here, replacing this text.

Question 3

Write a function that takes a filename as input and returns the sum of the numbers in the file. The file contains one number per line. The file is a text file with the extension .txt. The file is located in the same directory as the program. The file contains at most 1000 numbers. The numbers are integers and are between -1000 and 1000.

See the file `input_example.txt` for an example of the input file.

```
[ ]  def sum_all_from_file(filename: str) -> int:
         ...
```

```
[ ]  grader.check("q3")
```

Course metrics. A set of indicators was collected for both iterations of the course. These included measures of workload and turnaround (average grading time per problem, time from submission to grade release), student performance (average assignment grades, percentage of required resubmissions), and overall course outcomes (number of submissions, course evaluation scores). Descriptive comparisons across the two years were used to assess changes following the introduction of auto-grading.

Student feedback. Feedback from students was gathered on three official occasions: (i) verbally during the mid-course meeting with student representatives, held after the deadline of the first two assignments; (ii) in writing through the end-of-course evaluation survey; (iii) verbally during the course board meeting with student representatives, held several months after the conclusion of the course. Students were informed that automated feedback had been introduced for the first time in their iteration of the course. Students were also informed that anonymized comments from (i) and (iii) might be used for research and reporting purposes. Participation levels varied across the three occasions: in (i), five out of six student representatives attended; in (ii), the survey response rate was 10.94% (21 out of 192 students); and in (iii), only two student representatives were present.

Teaching assistant feedback. The two teaching assistants involved in the course were also invited to reflect on their experiences with the intervention. They were explicitly informed that their reflections could be included in this report and agreed to share their feedback for that purpose. Their reflections were provided by email correspondence.

4 Results

The introduction of auto-grading yielded measurable improvements in grading efficiency, feedback timeliness, and overall user satisfaction for both students and teaching assistants. Quantitative metrics and feedback from students and TAs are reported in what follows.

4.1 Course metrics

Table 1 summarizes the course metrics before and after the introduction of auto-grading. The clearest effect was a reduction in grading time for teaching assistants: in 2023, grading a single programming problem required on average about 5 minutes, whereas in 2024 this dropped to around 3 minutes. This represents roughly a 40% improvement in grading efficiency, freeing up time for other teaching activities such as student support.

Table 1: Key metrics before (2023) and after (2024) the introduction of auto-grading

Metrics	2023	2024
Number of enrolled students	168	193
Number of submissions - Assignment 1	157	180
Number of submissions - Assignment 2	148	173
Number of resubmissions required*	15%	5%
Average grade (pts. %)	76.6%	75.2%
Average grading time per problem (minutes)	$\sim$5	$\sim$3
Time from submission deadline to grade release (days)	14	9
Course evaluation - Assessment	3.9/5	4.0/5
Course evaluation - Overall impression	3.1/5	3.6/5

*Percentage of students required to resubmit at least one assignment due to insufficient or incorrect initial submission.

Another notable change was the decrease in resubmissions. In 2023, 15% of students had to resubmit at least one assignment due to errors or incomplete work; in 2024 this proportion fell to 5%. This indicates that immediate feedback from the auto-grading system helped students detect and correct mistakes earlier, leading to higher-quality submissions on the first attempt.

Average assignment grades remained stable (76.6% in 2023 vs. 75.2% in 2024), with no statistically significant difference. This suggests that the redesigned assignments did not make the tasks easier, but rather allowed students to reach similar outcomes more efficiently.

Turnaround time between submission and release of grades also improved. Students in 2023 typically waited about 14 days for results, compared with 9 days in 2024.

Finally, the overall course evaluations improved following the intervention. Student ratings for the assessment component rose slightly (from 3.9/5 in 2023 to 4.0/5 in 2024), while the overall impression of the course increased more substantially (from 3.1/5 to 3.6/5). These scores suggest that the redesign not only improved efficiency but also contributed positively to students' perception of the course experience.

Taken together, these results point to clear operational benefits for teaching staff and a more positive learning experience for students, providing the context for the feedback reported in the following subsections.

4.2 Feedback from students

In what follows, only comments specifically related to the use of auto-grading are reported. All such comments are included, grouped thematically into three main areas: perceived learning benefits, feedback clarity, and technical aspects.

Perceived learning benefits. Several students highlighted the value of immediate automated feedback for supporting their learning. They described how the system increased their confidence and helped them identify mistakes earlier:

> *"I really like the automated feedback with Otter-Grader! It guides me towards the solution and gives me confidence that I am on the right track." (i)*

> *"The automated feedback improves my learning journey because it is immediate and the results of the automated tests helps me understanding where I made mistakes." (i)*

The inclusion of Assignment 0 was also appreciated as a low-stakes way to familiarize themselves with the tool:

> *"Nice to have Assignment n.0 to test the system." (i)*

Other students valued the ability to run tests directly and to monitor their progress:

> *"I appreciated the otter-grader function on which you could test your program directly." (ii)*

> *"I feel like the first half of the course was pretty good, with the Jupyter Notebook submissions that had these self-graders that gave a nice indication if you were on the right track or not." (ii)*

One student even emphasized that Otter-Grader should remain in the course design:

> *"What should be kept for the next round of this course? Otter grader." (ii)* and *(iii)*

Feedback clarity. Although the overall experience was positive, some students noted that feedback could be more transparent. In particular, the use of hidden tests was occasionally perceived as limiting:

> *"Good to have the expected results in the assignments. Sometimes they could be more explicit, instead of being hidden in the tests of Otter-Grader." (i)*

Technical aspects. A few students mentioned practical difficulties related to the installation of the system, pointing to the need for continued technical support:

> *"It was not my case, but I know that someone struggled to install Otter-Grader and get it working." (i)*

4.3 Feedback from teaching assistants

Two teaching assistants contributed reflections. TA-1 had experience with both the 2022/23 (manual grading) and 2023/24 (partial auto-grading) iterations of the course, while TA-2 was only involved in the latter. Their comments focused on three main areas: workload and efficiency, learning curve, and quality of feedback.

Workload and efficiency. Both TAs emphasized that Otter-Grader reduced the burden of repetitive manual corrections. TA-1 described the previous year as *"extremely challenging, requiring extensive time and effort for assignment corrections,"* whereas the new setup *"significantly eased the workload despite minor issues."* Similarly, TA-2 noted that the tool helped avoid tedious manual work and reduced the risk of errors in grading.

Learning curve. TA-1 reported that Otter-Grader initially had a steep learning curve, but once mastered, it simplified the correction process: *"Although initially Otter-Grader had a steep learning curve, it makes the corrections much easier than just manual corrections (even when it fails!)."*

Quality of feedback. The system enabled more systematic detection of errors through batch evaluation and checkpoints, while also freeing up time for richer feedback. As TA-2 put it, *"It also frees up time to give more in-depth constructive feedback to students."* At the same time, TA-1 stressed the continued need to manually open notebooks to ensure students were not bypassing the automated checks.

In summary, the teaching assistants agreed that the integration of Otter-Grader made grading more efficient and consistent, while still requiring complementary manual review.

5 Discussion and Conclusions

The intervention was designed to address the research questions by examining changes in efficiency (RQ1), submission quality and performance (RQ2), and user experience (RQ3). The findings show that partial auto-grading with immediate feedback offered both pedagogical and operational benefits in the context of a large undergraduate AI course.

Impact on learning and teaching practice

The results for RQ1–RQ3 suggest that the intervention supported student learning while also improving the grading process. Students reported that immediate feedback increased their confidence, helped them identify mistakes earlier, and enabled higher-quality submissions on the first attempt, as reflected in the lower resubmission rate. At the same time, average grades remained stable, indicating that the assignments were not made easier, but that students were able to reach comparable outcomes more efficiently. Design choices such as including a preliminary *Assignment 0* further supported onboarding and reduced technical barriers.

From the teaching perspective, grading became faster and more consistent, allowing assistants to dedicate more time to higher-value activities such as feedback and support. Reflections from TAs confirmed that Otter-Grader eased workload and provided useful tools (e.g., batch evaluation and checkpoints), though manual review was still necessary to ensure reliability. Importantly, the introduction of auto-grading also required some additional redesign effort: compared with the usual yearly updates of assignments, instructors had to plan and validate automated tests in addition to revising the tasks themselves. This added workload was moderate, however, and considered manageable within the normal course preparation cycle, especially given the resulting gains in grading efficiency and feedback quality. Together, these results indicate that partial auto-grading can enhance both the pedagogical experience for students and the operational sustainability of large-enrollment courses.

Limitations and Challenges

Despite these positive outcomes, some limitations were noted. A few students experienced difficulties installing Otter-Grader on their local machines, highlighting the variability of student computing environments. Although these challenges are typical of software-based interventions, they underscore the need for robust technical support and clear documentation. In response, the 2025 course iteration introduced a cloud-based alternative (Google Colab with pre-installed tools) to mitigate installation issues—an important step toward greater accessibility.

Another concern, also highlighted in the literature, is the risk that students may become overly reliant on automated feedback, potentially hindering the development of essential skills such as independently testing and debugging their code. This limitation can be addressed through careful assignment design—for example, by withholding automated feedback for certain parts of the code. Such an approach encourages students to engage more deeply with the problem-solving process and to practice verifying their solutions without relying solely on external validation.

Another limitation relates to the scope of automated evaluation. While Otter-Grader provides valuable support for functional correctness, it does not fully address more nuanced aspects of programming such as code efficiency, readability, or testing strategy. To avoid misinterpretations, students were explicitly informed that the automated feedback was not equivalent to a final grade but rather a guide to support the problem-solving process.

Conclusion and Future Work

This intervention illustrates that a partial implementation of auto-grading can significantly enhance the effectiveness and efficiency of coding assignments in higher education. The approach supports formative learning, reduces grading effort, and maintains pedagogical integrity when paired with clear guidance and complementary manual review.

Future improvements could include expanding the range of exercises that use automated feedback, refining feedback messages to promote deeper reflection, and further simplifying the technical setup for students.

For educators considering similar implementations, this case highlights the importance of balancing automation with human oversight, designing user-friendly onboarding processes, and clearly communicating the role of automated tools within the broader assessment strategy.

References

Aldriye, H., Alkhalaf, A., & Alkhalaf, M. (2019). Automated grading systems for programming assignments: A literature review. *International Journal of Advanced Computer Science and Applications*, *10*(3). doi: 10.14569/IJACSA.2019.0100328

Douce, C., Livingstone, D., & Orwell, J. (2005, sep). Automatic test-based assessment of programming: A review. *J. Educ. Resour. Comput.*, *5*(3), 4-es. doi: 10.1145/1163405.1163409

Edwards, S. H., & Pérez-Quiñones, M. A. (2008). Web-cat: automatically grading programming assignments. In *Annual conference on innovation and technology in computer science education.* doi: 10.1145/1597849.1384371

Hegarty-Kelly, E., & Mooney, D. A. (2021). Analysis of an automatic grading system within first year computer science programming modules. In *Proceedings of the 5th conference on computing education practice* (p. 17-20). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3437914.3437973

Mitra, J. (2023). Studying the impact of auto-graders giving immediate feedback in programming assignments. In *Proceedings of the 54th acm technical symposium on computer science education v. 1* (p. 388-394). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3545945.3569726

Nikhila, K. N., & Chakrabarti, S. K. (2022). Letgrade: An automated grading system for programming assignments. In M. M. Rodrigo, N. Matsuda, A. I. Cristea, & V. Dimitrova (Eds.), *Artificial intelligence in education. posters and late breaking results, workshops and tutorials, industry and innovation tracks, practitioners' and doctoral consortium* (pp. 383–386). Cham: Springer International Publishing.

Paiva, J. C., Leal, J. P., & Figueira, A. (2022, jun). Automated assessment in computer science education: A state-of-the-art review. *ACM Trans. Comput. Educ.*, *22*(3). doi: 10.1145/3513140

Pyles, C., & UC Berkeley Data Science Education Program. (2024, March). *Otter-Grader: A Python and R autograding solution.* doi: 10.5281/zenodo.5259955

Participation in close-to-practice educational research - researchers' perspective

Samuel Bengmark
samuel@chalmers.se

Philip Gerlee
gerlee@chalmers.se

Jonathan Weidow
jonathan.weidow@chalmers.se

August 25, 2025

Sammanfattning

This study explores how researchers at a technical university experience and evaluate their participation in close-to-practice educational research within the framework of the Swedish ULF initiative. Drawing on ten semi-structured interviews analyzed using reflexive thematic analysis, the study identifies three primary domains of perceived value: career benefits, intellectual and professional stimulation, and meaningful collaboration with schools. A fourth theme highlights the enabling or constraining role of institutional and project-related conditions, such as time allocation and project clarity. While early-career researchers viewed close-to-practice participation as strategically beneficial, more senior researchers emphasized its contributions to learning, professional identity, and societal engagement. The findings suggest that the value of close-to-practice research is constructed through both personal and institutional contexts, and that successful participation depends on individual motivation as well as structural support. This study contributes to a deeper understanding of researcher engagement in practice-oriented collaborations and offers insights into how partnerships between universities and schools can be more effectively designed and supported.

Sammanfattning

Denna studie undersöker hur forskare vid ett tekniskt universitet upplever och värderar sitt deltagande i praktiknära utbildningsforskning inom ramen för det svenska ULF-initiativet. Baserat på tio semistrukturerade intervjuer och analyserade med hjälp av reflexiv tematisk analys identifierar studien tre huvudsakliga domäner av upplevt värde: karriärfördelar, intellektuell och professionell stimulans samt meningsfullt inverkan och samverkan. Ett fjärde tema belyser den möjliggörande eller begränsande roll som institutionella och projektrelaterade villkor, såsom tidsfördelning och projektets tydlighet, spelar. Medan forskare i början av sin karriär såg deltagande i praktiknära forskning som strategiskt fördelaktigt, betonade mer seniora forskare dess bidrag till lärande, professionell identitet och samhälleligt engagemang. Resultaten tyder på att värdet av praktiknära forskning konstrueras i relation till både personliga och institutionella sammanhang, och att framgångsrikt deltagande beror på såväl individuell motivation som strukturellt stöd. Studien bidrar till en fördjupad förståelse av forskares engagemang i praktiknära samarbeten och erbjuder insikter i hur partnerskap mellan universitet och skolor kan utformas och stödjas mer effektivt.

Keywords: *close-to-practice research; technical university; value;*

1 Introduction

In recent years, there has been growing interest in close-to-practice (CtP) educational research as a means of bridging the gap between academic inquiry and the practical realities of teaching and learning (BERA-RSA, 2014; Parsons, 2021). CtP research is typically conducted in authentic educational settings and is motivated by the needs, questions, and challenges experienced by practitioners. It seeks to generate contextually relevant knowledge that can inform both practice and theory, often through collaboration between researchers and educators (Hordern, 2021; Mawson, Serret, Woodford, Greeley, & Lindsell, 2024)

The Swedish ULF initiative, short for "Utbildning, Lärande, Forskning" (Education, Learning, Research), is a national effort to promote CtP research by fostering long-term collaboration between universities and schools. Since 2018, Chalmers University of Technology has participated in this initiative, involving researchers in technical disciplines in collaborative projects with upper secondary schools. These projects typically revolve around jointly developed inquiries, sometimes supported by student thesis work, and aim to generate both didactic insights and practical innovations relevant to STEM education.

Although much of the existing literature on CtP research has focused on teacher experiences or organizational structures that allow such collaborations (Lewis, Brooks, Parker, & Thomas, 2018; Parsons, 2021), less attention has been paid to how researchers, particularly those from non-education disciplines, experience and make sense of their participation in such initiatives. Understanding their perspectives is critical for improving the design and sustainability of university-school partnerships and for fostering a research culture that genuinely values practice-based knowledge production.

While the number of participants is limited, this study explores how researchers at Chalmers perceive the value of their engagement in CtP research projects in search of answers to the following research question:

What do researchers at a technical university consider valuable in participating in close-to-practice educational research?

2 The local context - ULF at Chalmers

At Chalmers University of Technology there has been CtP research activities within the ULF-framework since 2018. It is organized according to an academic year structure. The process begins in mid-spring, when Chalmers' partner schools submit project proposals. These projects are then evaluated based on three main criteria: feasibility, added value to Swedish schools, and the availability of a suitable researcher at Chalmers. During this phase, a dialogue is also held with the individuals who submitted the proposals. At the end of spring, the selected projects are announced, and an agreement is signed between Chalmers and the participating school.

During the autumn term, collaboration begins between the Chalmers researcher and the participating upper secondary school teachers in order to further shape the project. A thesis proposal is developed and made visible to Chalmers students by mid-autumn. If the project attracts thesis students, they will contribute many hundreds of hours of work to the project during the spring. In addition to the thesis report, it is desirable that the researcher and the teachers jointly author a report to be published in a suitable forum.

The expectations on the researcher from Chalmers is that they will provide i) general scientific expertise, e.g. developing the teachers ideas into a concrete research question, ii) domain knowledge in the relevant subject e.g. mathematics and iii) if possible provide a suitable theoretical grounding of the research question in didactical theory. The researcher will typically receive appproximately 100 hour working hours to contribute to the project during one academic year.

Participation in practice-based research offers several valuable benefits for Chalmers. The participating researchers develop their competence in didactic research, which if properly utilized—can help enhance the quality of undergraduate education at Chalmers. Furthermore, publishing papers in didactical research contributes to raising the visibility of Chalmers' expertise in this area. The government has also highlighted practice-based research as a means to address a societal challenge—the quality of education for future generations. Improving the Swedish school system gives Chalmers yet another pathway for contributing societal value, both through research expertise applied to improve schools and, over time, through the practice-based research findings themselves.

The broader community also benefits greatly from this type of research. Since practice-based research is rooted in the needs of schools, Chalmers can help strengthen the quality and evidence base of the education system. The opportunity for teachers to collaborate with the university also offers mutual professional development and stimulating tasks, which may help make the teaching profession more attractive. Also, an improvement in the quality of Chalmers' partner schools hopefully leads to an improvement in the quality of Chalmers' teacher education programs.

However, there are challenges. Organising ULF we have noted a difficulty to attract researchers at Chalmers. Anecdotal evidence suggests that this is due to limited knowledge of ULF and lack of time on the part of the researchers. In order to improve practice-based research at Chalmers we therefore decided to study participation in ULF in a systematic manner.

3 Methodology

This study employs a qualitative research design informed by a reflexive thematic analysis approach by Braun and Clarke (2006, 2019) to explore how researchers at Chalmers perceive the value of participating in CtP educational research. This offers a flexible yet rigorous method for identifying, analyzing, and interpreting patterns of meaning across qualitative data. The analysis was inductive, meaning that themes were generated from the data. The aim is to gain an in-depth understanding of the experiences, motivations, and perceived outcomes among academics involved in ULF projects—a national initiative designed to foster collaboration between universities and schools through practice-oriented research.

The empirical material consists of semi-structured interviews with 10 researchers who had been involved in CtP research at Chalmers. In total 13 researchers have participated since the start at Chalmers in 2018, of which three are authors of this article. A majority of the researchers have a recorded previous interest in didactical research (e.g. publications), but only a few have formal training in didactics.

The interviews were guided by an interview protocol, designed to elicit reflections on motivations for participating, experiences of collaboration, use of research time, perceived outcomes, and views on future involvement. Interviews were conducted between March and April 2025 and ranged from approximately 30 to 60 minutes in duration. All interviews were conducted by the authors in pairs and were audio recorded and transcribed verbatim. All interview transcripts were read in full to gain a holistic sense of the material. Initial notes were made on recurring concepts or striking expressions related to value, motivation, and collaboration.

Segments of text that captured explicit or implicit references to perceived value, including references to career development, contribution to practice, learning, institutional support, and collaboration were identified. These were grouped into themes based on conceptual similarity. At this stage, potential overarching themes began to emerge.

Themes were reviewed against the full dataset to ensure internal consistency and distinctiveness. Some candidate themes were merged, split, or redefined to better capture the complexity of participants' perspectives. Each theme was given a clear conceptual definition. Illustrative quotes were selected to exemplify key aspects of the themes.

All participants gave informed consent. The project adheres to the ethical guidelines of the Swedish Research Council. Care has been taken to ensure that participants cannot be identified in the published analysis.

4 Results

Three overarching themes were identified: *(1) Career benefits*, *(2) Intellectual and professional stimulation*, and *(3) Meaningful contribution and collaboration*. A fourth cross-cutting theme, *(4) Conditions enabling engagement*, emerged as a moderating factor influencing the perceived value of participation.

4.1 Career benefits

Some researchers expected that the participation in the ULF-project would benefit their career. Characteristic for this group was that they were all early-career educational researchers aiming at publishing academic journal articles about this CtP research project.

> *"It was a great step toward broadening my research... very beneficial for my academic track record."*

One of them had already seen the positive effect in a hiring process.

> *"It felt like a good opportunity... something that I truly believe was also a merit when I later applied for and got this position."*

The more senior researchers did not expect that participating in an ULF-project would affect their career or qualifications. For some of them, the project was never completed and the others did not think that the project was good enough fit with their interests. One senior respondent considered the project very interesting and in line with their research but that it came too late in their career.

4.2 Intellectual and professional stimulation

Several participants emphasized that working closely with school practitioners and in real-world classroom contexts offered a refreshing and intellectually engaging experience. Several described the complexity of classroom environments and the process of co-constructing a research project as both demanding and invigorating.

> *"It was a step outside my comfort zone... but also very educational to take on pedagogical questions."*

One of the younger researchers explicitly identified project leadership as a positive and growth-oriented experience within the ULF context.

> *"I've been part of primary projects before... but only as a contributor, but now I get to be the project leader for this project. I'm learning a lot from planning an entire project from start to finish. It's very educational for me as a researcher."*

4.3 Meaningful impact and collaboration

Engagement in ULF projects was often described as a way of contributing to practice, to society — not just extracting data.

> *"It's an incredibly important contribution to the society around us... I would really like to see us find ways to contribute. This is a very concrete and positive way to do it*

> *"Maybe we at Chalmers could make significant contributions in this area."*

There was a wish among several researchers to be involved in activities that affect school education.

> *"Contribute to strengthening teaching at other levels as well... a service to society in a certain sense."*

> *"What I have seen in this teaching also works in other classroom situations."*

Participants valued the opportunity to work closely with teachers, experiencing mutual respect and co-learning. The collaboration with the school teachers was stimulating.

> *"It was a very rewarding experience... they tested everything I had worked on in practice."*

Some expressed that they were interested in learning more about school or this specific school level.

> *"I was curious about how upper secondary school teachers think and what their needs are."*

> *"For me, it was actually a fun way to get more involved in upper secondary school activities."*

Moreover, collaborative dynamics between researchers and teachers sometimes created environments of trust that enhanced motivation and enjoyment.

4.4 Conditions enabling engagement

This cross-cutting theme captured factors that either enhanced or hindered the sense of value researchers derived. Chief among these were *time allocation*, *project clarity*, and *institutional support*.

Some researchers reported having ample time and flexibility as the ULF-project was integrated in their regular research time, which allowed them to engage deeply:

> *"[Research is] pretty much my entire position. That's why I applied for this."*

Others described more constrained conditions:

> *"The ULF project was counted as teaching time for me... it became departmental service."*

Lacking time to write down the results from the project was a something several participants mentioned. This hindered some from pushing the project all the way.

"The project never really happened... because other things became more important."

Some projects managed to attract a master's student from Chalmers who carried out his or her master's thesis. This was considered advantageous, as it enabled more total hours to be spent on the project.

The clarity of project goals at initiation also played a key role. Projects with vague or evolving aims sometimes led to frustration:

"The project itself was poorly defined from the start... a clear disadvantage."

In these cases it was also coupled with bad matching, such that the researcher did not feel at home with the methods needed in project.

Conversely, others found that the ULF-project was already set up, with authentic questions from the practice and domains for data collection. This made it easy to slip into the project.

5 Discussion and conclusions

For early-career educational researchers, involvement in ULF projects offered tangible career advantages such as CV-building and leadership opportunities. This resonates with Lewis et al. (2018), who found that CtP research can help researchers develop a stronger professional identity. For senior researchers the academic payoff was often perceived as limited. This discrepancy points to the potential importance of career stage, or incentive mechanisms, when involving researchers in CtP research.

Some participants described their engagement in terms of professional renewal and intellectual curiosity. The opportunity to co-develop research with practitioners, grapple with didactical challenges, and operate outside their disciplinary comfort zones was described as both stimulating and educational. This view is in agreement with Parsons' (2021) emphasis on the potential of collaborative spaces that bridge academic and professional knowledge.

Participants also viewed their involvement as a form of civic contribution, aligning with Mawson et al.'s (2024) argument that CtP research can foster a sense of purpose through local relevance and social impact.

The researchers' ability to realize these forms of value was conditioned by structural factors. Adequate time allocation, institutional support, and prepared project emerged as critical enablers. In contrast, vague project scopes or mismatches between research expertise and project needs led to frustration. This echoes Hordern's (2020) critique that when practice is often organizationally under-supported.

Taken together, our findings suggest that CtP research holds significant value for researchers at technical universities. However, success depends on thoughtful design, alignment with academic identities, and institutional scaffolding. It is not merely proximity to practice that matters, but the conditions under which collaboration is initiated and sustained.

Our findings suggest three areas for improvement to focus on. Firstly, that CtP research should be better integrated with researchers' existing scholarly agendas, particularly for senior staff. Secondly, early-career researchers may benefit from targeted support and recognition for their leadership in such projects. Finally, there is a need for structures that ensure clarity of purpose, access to data, and time for dissemination. Together, these can significantly improve the perceived value of participation.

5.1 Limitations and future research

This study is limited by its scope to one technical university and a relatively small sample. Future research could examine how these patterns compare across other disciplines or institutional types. Longitudinal studies might also explore whether and how participation in CtP research affects academic careers and teaching practice over time.

References

BERA-RSA. (2014). *The role of research in teacher education: Reviewing the evidence.* I.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. doi: 10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597. doi: 10.1080/2159676X.2019.1628806

Hordern, J. (2021). Why close to practice is not enough: Neglecting practice in educational research. *British Educational Research Journal*, *47*(6), 1451–1465.

Lewis, H., Brooks, S., Parker, G., & Thomas, D. (2018). A close-up on close-to-practice research: Reflecting on teacher educators' experiences of and engagement with a classroom-based research project. *Wales Journal of Education*, *20*(1).

Mawson, K., Serret, N., Woodford, M., Greeley, L., & Lindsell, H. (2024). Professional development through close to practice partnerships between teachers and researchers. In *Iceri2024 proceedings* (pp. 6184–6184).

Parsons, S. (2021). The importance of collaboration for knowledge co-construction in 'close-to-practice'research. *British Educational Research Journal*, *47*(6), 1490–1499.

Work in progress

Swinging Pendulums on the Cloud: Digitization of Simulation & Experimental Infrastructure for Feedback-based Active Learning

Shivesh Kumar
Dynamics Division, Department of Mechanics & Maritime Sciences,
Chalmers University of Technology, Gothenburg, Sweden
shivesh.kumar@chalmers.se

January 16, 2026

Abstract

This paper presents the CloudPendulum platform - an open-access, cloud-based system designed to enhance robotics and AI education through interactive programming coursework. CloudPendulum enables students and instructors to engage with both dynamic simulations and live experimentation hardware (such as simple pendulum systems) remotely, fostering active and experiential learning. We deployed the platform in two pedagogical scenarios at Chalmers: one where only the instructor accessed the hardware, and another where both students and the instructor had direct access. Results indicate that student access to the hardware enabled by our cloud technology led to higher session ratings in terms of overall quality, instructional clarity, and engagement. Students also expressed strong interest in broader adoption of the platform in future courses.

Abstract

Denna artikel presenterar plattformen CloudPendulum - ett öppet, molnbaserat system som är utformat för att förbättra robotik- och AI-utbildningen genom interaktiva programmeringskurser. CloudPendulum gör det möjligt för studenter och lärare att på distans interagera med både dynamiska simuleringar och hårdvara för liveexperiment (såsom enkla pendelsystem), vilket främjar aktivt och erfarenhetsbaserat lärande. Vi implementerade plattformen i två pedagogiska scenarier vid Chalmers: ett där endast läraren hade tillgång till hårdvaran och ett där både studenter och lärare hade direkt tillgång. Resultaten visar att studenternas tillgång till hårdvaran, möjliggjord av vår molnteknik, ledde till högre betyg för sessionerna när det gäller övergripande kvalitet, tydlighet i undervisningen och engagemang. Studenterna uttryckte också ett stort intresse för en bredare användning av plattformen i framtida kurser.

Keywords: *cloud robotics; STEM education; active learning; remote labs.*

1 Introduction

Traditional approaches to teaching robotics theory often rely on mathematical derivations, manual problem-solving, and simulations implemented in proprietary or complex software environments. While these methods have proven effective in foundational STEM education, they struggle to meet the demands of a rapidly evolving robotics landscape, where systems are increasingly complex and interdisciplinary as discussed in Ryalat, Almtireen, Al-refai, Elmoaqet, and Rawashdeh (2025). To foster deeper understanding, it is essential that students validate mathematical models through both hardware experimentation and physics-based simulations. Without this, learners may fail to grasp the scope and limitations of their models according to Choi et al. (2021). However, access to physical robot hardware remains a significant barrier due to cost, fragility, and maintenance requirements. Similarly, simulation environments, though valuable, are often platform-dependent, difficult to install, and lack well-maintained open-source alternatives. Recent advances in cloud computing and Internet of Things (IoT) technologies offer promising solutions. Platforms such as JupyterHub, Google Colab, and Deepnote enable students to interact with simulations and even remote hardware through browser-based interfaces, eliminating the need for specialized local setups. These tools support real-time collaboration, reproducibility, and accessibility across devices, making robotics and AI education more inclusive and scalable. By leveraging these technologies, educators can design learning experiences that combine theoretical rigor with hands-on experimentation - aligning with active learning principles Freeman et al. (2014) which has shown to significantly improve student outcomes in STEM disciplines. This approach can also trigger higher-order learning outcomes such as apply, analyze, evaluate, and create in Bloom's taxonomy Bloom et al. (1956), especially for students who may not have access to expensive robotics infrastructure. This shift has the potential to democratize robotics education, allowing students to engage with real-world systems from anywhere, using only an internet-enabled device.

The aim of this paper is to introduce and evaluate a platform-independent framework for robotics and AI education accessible on any target computer with a web browser, regardless of hardware architecture, operating system, or proprietary software requirements. The deployment of the system in a web-based environment (e.g., Jupyter Hub, Google Colab or Deepnote) is explored, thus not requiring students to install software on their own computer, which is always a difficult step given the many different platforms used by students (Windows/Linux/Mac). Using the proposed technology, some tutorials and exercise sessions are developed, using which students can learn basic concepts in dynamics and control (e.g., forward and inverse dynamics and their application in simulation and gravity compensation, energy shaping and LQR control) on simple canonical mechanical systems (e.g. simple pendulum). To judge the technological readiness and its impact on pedagogy, two pilot studies are performed in two different courses namely MMA093 - Rigid Body Dynamics course in Study Period 2 and Tracks course TRA455 on Athletic Intelligence in Robotics (AIR) in Study Period 4 in the academic year 2024-2025.

2 Related Work

Active learning has emerged as a powerful pedagogical approach that improves student performance in STEM disciplines by engaging them in the learning process through activities that promote analysis, synthesis, and evaluation of class content as discussed in Freeman et al. (2014). One widely adopted model is the flipped classroom, where students first encounter new material outside of class - typically through readings or video lectures - and then use classroom time for interactive problem-solving and discussion, see Bishop and Verleger (2013). This model shifts the instructor's role from a lecturer to a facilitator of learning. Within this framework, Think-Pair-Share is a structured active learning technique where students first think individually about a question, then discuss their thoughts with a peer, and finally share their conclusions with the class. Owens and Tanner (2017) argue that this concept is successful for active learning because it involves the release of several neurotransmitters which boost synaptic plasticity which is a biological process associated with behavioral learning and memory formation.

In STEM education, integrating simulation and hardware platforms into teaching can further enhance active learning by enabling hands-on experimentation and collaborative exploration. Remote-controlled laboratories, in particular, offer new opportunities for such engagement. These labs leverage cloud computing, IoT, and AI to provide students with real-time access to physical systems from anywhere in the world. For a comprehensive overview, see Heradio, de la Torre, Galan, et al. (2016), and for their specific impact on control education, see Heradio, de la Torre, and Dormido (2016). Several recent initiatives exemplify this trend. The RealAIGym platform by Wiebe, Vyas, Maywald, Kumar, and Kirchner (2022) provides open source hardware and software design of canonical underactuated robotic systems—such as torque-limited pendulums Wiebe, Babel, et al. (2022), double pendulums by Wiebe et al. (2024), brachiating robots by Javadi et al. (2023), Grama, Javadi, Kumar, Boroujeni, and Kirchner (2024) and hopping legs by Albracht, Kumar, Vyas, and Kirchner (2024); Soni et al. (2023)—for benchmarking learning and control algorithms. However, their original system design is monolithic and bulky. For example, the RealAIGym's torque limited simple pendulum has a moving inertia of 0.125 kg m^2 and given maximum speed of motors is 40 rad/s, the system can store rotational kinetic energy of up to 100 Joules. This amount of mechanical energy raises safety concerns for continuous unsupervised remote access. In contrast, our work emphasizes modularity and safety to support 24x7 unsupervised use. Other notable efforts include CloudGripper by Zahid and Pokorny (2024) for collecting manipulation data at scale, Remote Labs at TU Eindhoven by van den Beemt, Groothuijsen, Özkan, and Hendrix (2023) for control theory education, and Drone Studio by Arizona State University Engineering (2025) which enables remote drone experimentation for research and education.

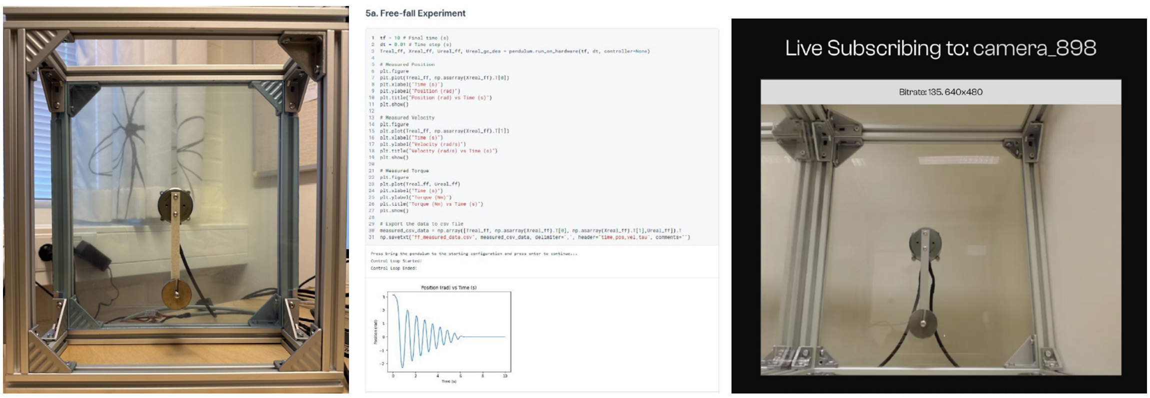


Figure 1: Swinging pendulums on the cloud project: a torque limited simple pendulum unit (left), Python based Jupyter notebook for remote experimentation (center), live camera feed accessible on a web-browser (right)

3 CloudPendulum Technology

A miniaturized torque-limited simple pendulum prototype inspired from the open source torque limited pendulum by Wiebe, Babel, et al. (2022) has been developed in this project. A mass of 50 g is connected on a rod of length 10 cm connected to a direct drive actuator for low friction and high mechanical transparency as shown in Figure 1 (left). The system has a rotational inertia of 5×10^{-6} kg m^2 and with the maximum rotational speed of 85 rad/s, the system can store a maximum rotational energy of 0.036 Joules which is orders of magnitude lower than the RealAIGym's torque limited simple pendulum making it much safer by design for remote unsupervised use. This simple mechanical system is connected to a desktop server computer via USB2CAN converter. The computer runs a Jupyter Hub server which makes it possible to control the system via a website: `cloudpendulum.m2.chalmers.se`. New users can create an account using the signup feature which is then validated by the system administrator to authorize their access to the remotely hosted hardware. The system provides a simulation and live camera stream (under 500 ms latency) from the real experimental hardware as well as position, velocity and torque measurements from the system to the users so that they can immediately see how their model performs on the real system. Currently, 3 identical pendulum units have been built and integrated in the CloudPendulum infrastructure which can be used by 20-25 users in parallel (subjected to worst case waiting time in order of a few seconds). Further, safety limits at position, velocity and torque levels have been implemented so that the system is protected from malicious commands that may affect system's health. This allows for safe unsupervised access to the system. By providing a scalable and 24x7 accessible environment, the platform is designed to significantly enhance active learning experiences through prompt educational feedback and physically insightful interactions.

4 Pedagogical Developments using CloudPendulum

The cloud pendulum system can be used in several ways to support teaching and research in STEM education. Since the designed cloud pendulum hardware is portable and easy to carry, the system can be used as a teaching aid in classrooms by the instructor. This may be combined with allowing remote independent access for students during the in-class activity. Feedback is collected from students to evaluate their learning experience as well as the ease of working with the cloud pendulum infrastructure. In the scope of this work, we evaluate its pedagogical use case in two different settings.

4.1 Pilot Study 1 in MMA093: Instructor has physical access to cloud pendulum hardware during the in-class activity.

During study period 2 of the academic year 2024-2025, the pendulum hardware is evaluated as a teaching aid in the MMA093: Rigid Body Dynamics classroom to demonstrate real experiments to students and the data collected from the hardware is made available to students for further analysis and programming coursework. Using the proposed hardware test bench, a module for teaching forward and inverse dynamics concepts is prepared and evaluated using student feedback. Note that only the instructor has physical access to pendulum hardware during the in-class activity (see Fig.2, left). We follow a flipped classroom approach for this endeavor:

1. A lecture on the topic is made available to students prior to the in-person classroom session via Chalmers play (Link here).

2. During the classroom session, students first derive the equations of motion for a simple pendulum by pen and paper. Then, they are asked to implement a simulation of the pendulum free-fall using a Python Notebook hosted on deepnote available here. Next, they implement a gravity compensation controller and test it in the simulation. Finally, the instructor demonstrates the working of the controller directly on the real system as well as collect data from the real system to demonstrate the accuracy of the simulator.

3. As an after-class activity, the students are asked to compare and reflect upon differences between their simulation model and the behavior of the real system. The students are asked to develop a method to close the simulation to reality gap.

The students are provided with a template code so that they do not have to program everything from scratch. Additionally, deepnote is integrated with an AI based auto-complete feature which can help students with programming syntax. The final tasks are intentionally crafted to require skills and reasoning that go beyond what the most widely available AI models can currently achieve (as of the session date). We recognize that AI capabilities are rapidly advancing, so it is essential for educators to continuously adapt and increase the complexity of such tasks to ensure that students are challenged and can achieve the intended learning outcomes. The notebooks are equipped with various Think-Pair-Share (TPS) activities so that students can actively play with the simulator and discuss their insights with fellow students. Fig. 2 (right) shows an example of such an activity where students are asked to play with different initial conditions, time step size in simulation to identify the fixed points of the system and comment on numerical accuracy of different methods.

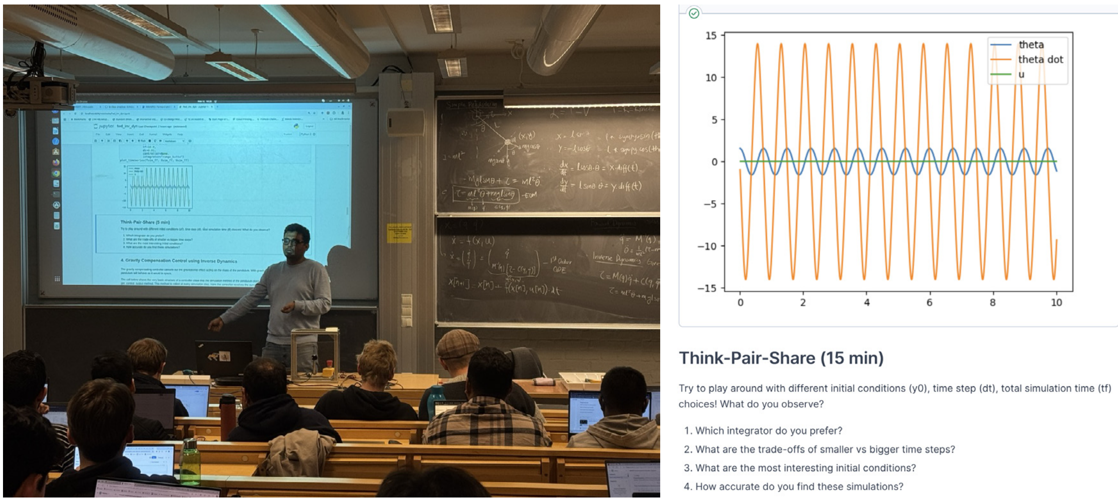


Figure 2: Impressions from pilot study 1 in MMA093 course: instructor with physical access to cloud pendulum hardware during the in-class activity (left), example of simulation based Think-Pair-Share activity (right)

4.2 Pilot Study 2 in TRA455: Instructor has physical access and students have remote access to pendulum hardware during the in-class activity.

During the study period 4 of the academic year 2024-2025, the cloud pendulum hardware was used in a tutorial session in TRA455: Athletic Intelligence in Robotics course in a modality where the instructor had physical access to the cloud pendulum hardware and the students are allowed remote access to the hardware for the in-class activity as shown in Fig. 3. The topic of the tutorial session consisted of Energy Shaping + Linear Quadratic Regulator (LQR) controls and iterative-LQR model predictive control for the swing-up task. We followed a Lecture-Tutorial-Practice (LTP) pedagogical method for this session.

1. Prior to the tutorial, a classical lecture was conducted to explain the trajectory stabilization methods (e.g. LQR, iLQR etc.) to students. After the lecture, the video recordings of the lecture was made available to students here.

2. During the tutorial session, two different Jupyter notebook environments on the topics Energy Shaping + LQR and iterative LQR MPC are provided to students via Chalmers Gitlab. Students are provided with user accounts on Cloud Pendulum JupyterHub so that they can connect to the remote hardware for experimentation during the session.

3. After the tutorial activity, the students were given a practice task where they derive an LQR controller for upright stabilization of a pendubot (double pendulum with shoulder actuation) and iterative LQR controller for swing-up and stabilization task for acrobot system (double pendulum with elbow actuation). Remote experimentation was not allowed for the practice session due to lack of double pendulum hardware hosted on the cloud.

As before, the students are provided with a template code so that they can focus on the ILOs. As opposed to the previous case, the students did not have access to AI autocomplete due to lack of integration with JupyterHub environment. The notebooks are similarly equipped with various TPS activities so that students can actively play with the simulator and remote hardware and discuss their insights with fellow students. As an example of TPS activity, the students were asked to tune the Q and R matrices for the LQR controller in simulation and hardware and observe the challenges in sim2real transfer.

Figure 3: Impressions from pilot study 2 in TRA455 course: instructor with physical access and students with remote access to cloud pendulum hardware during the in-class activity (left), 3 pendulum units connected to the cloud which students can connect to (right)

4.3 Data Collection and Analysis

The students were provided with an anonymous online survey at the end of the session to study their experience of participating in these unique in-class activities. The questionnaire included a total of 12 questions combining mandatory 5 point Likert scale type questions, multiple choice questions (MCQs) and optional free text comments. Attributes such as overall session content, clarity of instruction and engagement during in-class activity were measured on a scale of 1-5 stars. The student's grasp of the concrete ILOs in the session were measured using 5 point Likert scale ranging from *Not at all*, *Somewhat well*, *Well*, *Very Well* to *Extremely Well*. The students were asked to record their prior experience with the mathematics involved, Python programming and frameworks for hosting Python Notebooks (such as Deepnote or JupyterHub) with *Yes*, *No* and *Partly*. Free text questions included what they liked the most about the session, what could be improved in future and any additional comments or suggestions. Finally, the students were asked to record with a binary *Yes/No* question whether they liked the connection of mathematical modeling, simulation and hardware in a single interactive Python notebook and if they would like to see this technology adopted in their future courses at Chalmers. The survey was designed to be completed in 2–5 minutes of time and the online survey form may be accessed here.

5 Results & Discussion

To evaluate the pedagogical developments using CloudPendulum, the response to the online questionnaires given to the students attending the sessions were analyzed in both cases. Basic knowledge of Python and some associated libraries like numpy, sympy was needed to follow the content of the sessions as well as basic familiarity with Jupyter notebooks was necessary but was not communicated to the students in advance.

Participants A total of 50 students were enrolled in MMA093 course out of which around 30 attended the in-class session for the first pilot study. From this, 25 students answered the survey resulting in a participating rate of 83.3%. A total of 25 students were enrolled in the TRA455 course out of which around 9 attended the in-class session. From this, 9 students answered the survey resulting in a participating rate of 100%. In both cases, the attendance in these special sessions was not mandatory.

Quantitative Results Table 1 provides a comparative analysis of selected quantitative questions from the survey. It's evident that both pilot studies received generally positive ratings from the students. Metrics such as overall session quality, clarity of instruction and engagement during the session received slightly higher scores in pilot study 2 in TRA455 where both instructor as well as students had access to the pendulum hardware as compared to pilot study 1 in MMA093 where only the instructor had access to the pendulum hardware. This demonstrates that providing students with individual access to hardware via our proposed cloud platform can make classroom sessions more engaging for students. All participants liked the connection between mathematical modeling, simulation and the hardware in a single interactive Python notebook environment. For MMA093, 88% of the students voted for seeing this kind of technology adopted in their future courses at Chalmers. It must be noted that around 80% of the students had some background in Python and the students who gave a bit lower score could be correlated to the students who joined the session with no understanding of Python language. On the contrary, during the pilot study in the TRA455 course, 100% of the students voted for seeing this kind of technology adopted in their future courses at Chalmers. It must be noted that all of the students attending the tutorial session in TRA455 had some background in Python. Comparing the student feedback in the two pilot studies, it can be concluded that former understanding of Python and Jupyter notebooks enhances the student experience in such novel teaching scenarios. Hence, it is important that students and teachers using the platform in their education ensure that they have sufficient knowledge of Python. Also, some basic familiarity with Jupyter notebooks or Deep note interface is nice to have.

Qualitative Results Fig. 4 shows the free text responses to the survey question about what the students liked the most about the sessions. In the written feedback of MMA093 pilot study, the students appreciated the evident connection between theory and practice, its hands on nature, the obvious connection between mathematics and real life, and the interactive aspect. In the written feedback of TRA455 pilot study, the students highly appreciated the ability to perform experiments in simulation as well as real hardware using the cloud pendulum hardware. They particularly enjoyed viewing the live stream of how

Table 1: Comparative analysis of survey results in the two pilot studies

Survey Question	MMA093	TRA455
Overall Session Quality (mean $\pm$ std. dev., out of 5)	4.76 ± 0.43	5
Clarity of Instruction (mean $\pm$ std. dev., out of 5)	4.52 ± 0.58	4.89 ± 0.33
Engagement during the Session (mean $\pm$ std. dev., out of 5)	4.68 ± 0.55	4.78 ± 0.44
Students with Prior Background in Python	80%	100%
Students vouching for Increased Future Use of CP	88%	100%
Liked Integrated Modeling, Simulation and Hardware	100%	100%

their controller implementations were performing on the real system with minimum lag. Thanks to the smart scheduling system, none of the students complained about waiting times.

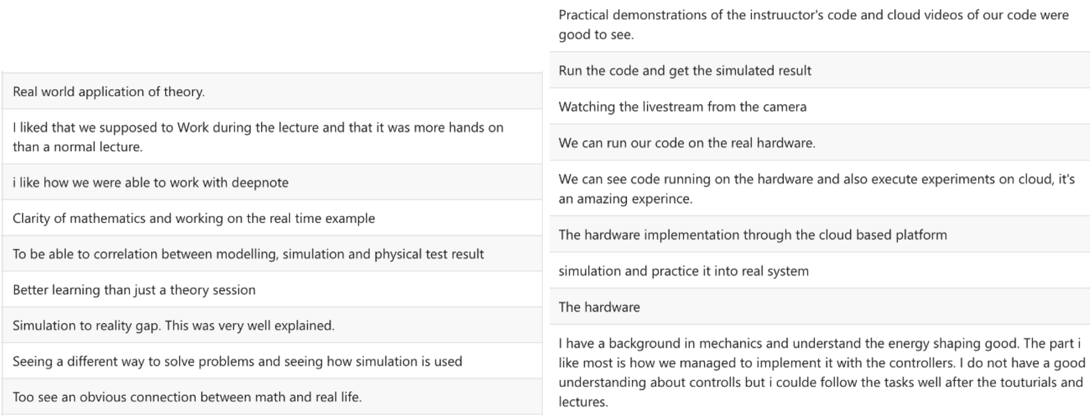


Figure 4: Responses to *What did you like most about the session?*: MMA093 Pilot Study (left), TRA455 Pilot Study (right)

Limitations The pilot studies reported here involved relatively small sample sizes, particularly in Pilot Study 2, which included only nine participants. As a result, the quantitative findings should be interpreted as indicative rather than conclusive, with limited statistical significance due to the absence of a control group. In MMA093, most students were enrolled in the Master's programs in Applied Mechanics and Mobility Engineering at Chalmers, and were unfamiliar with the instructor's teaching style. In contrast, students in TRA455—primarily from Systems & Control, Complex Adaptive Systems, and Applied Mechanics programs—had prior exposure to the instructor. This difference in familiarity may have introduced bias in student responses. Furthermore, the study captures only the short-term impressions of students immediately following the sessions. Longitudinal studies are needed to assess the sustained impact of integrating hardware-based experiences into robotics and AI education, particularly in terms of skill development, conceptual understanding, and long-term retention.

6 Conclusion and Outlook

In this paper, we have presented a novel CloudPendulum system which is a digital platform for programming coursework for students and teachers in robotics and AI education. This scalable and 24x7 accessible environment provides prompt educational and physically insightful feedback, supporting learning in both theoretical and practical aspects of their education. We have evaluated its use in two different pedagogical scenarios. It is shown that the system has the potential to make the classroom sessions more engaging and fun for students. The target groups are students and teachers of courses with programming assignments related to artificial intelligence, dynamics, robotics, and control. In future, we will design practice assignments for students where the students can use the cloud pendulum hardware to solve their homework tasks 24x7. Additionally, the usage of this system in benchmarking learning and control methods with RealAIGym inspired leaderboards will be explored.

Acknowledgements

The author would like to thank Damodar Datta Kancharla for assisting the early hardware and software development in this project and Joan Badia for testing the cloud pendulum system and tutorial development in the TRA455 course. The work presented here is supported by grants from the Chalmers Foundation Digitalization Initiative (Grant Number: T-2406-3947) and Chalmers Innovation Office (Grant Number: CIK-24-067).

References

Albracht, M., Kumar, S., Vyas, S., & Kirchner, F. (2024). Model predictive parkour control of a monoped hopper in dynamically changing environments. *IEEE Robotics and Automation Letters*, *9*(10), 8507-8514. doi: 10.1109/LRA.2024.3445668

Arizona State University Engineering. (2025). *Drone research at asu.* Online. Retrieved from `https://engineering.asu.edu/drone/` (Accessed: 2025-03-19)

Bishop, J., & Verleger, M. A. (2013). The flipped classroom: A survey of the research. In *2013 asee annual conference & exposition* (pp. 23–1200).

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., Krathwohl, D. R., et al. (1956). *Taxonomy of educational objectives: The classification of educational goals. handbook 1: Cognitive domain*. Longman New York.

Choi, H., Crump, C., Duriez, C., Elmquist, A., Hager, G., Han, D., ... Trinkle, J. (2021). On the use of simulation in robotics: Opportunities, challenges, and suggestions for moving forward. *Proceedings of the National Academy of Sciences*, *118*(1), e1907856118. Retrieved from `https://www.pnas.org/doi/abs/10.1073/pnas.1907856118` doi: 10.1073/pnas.1907856118

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the national academy of sciences*, *111*(23), 8410–8415.

Grama, S. S., Javadi, M., Kumar, S., Boroujeni, H. Z., & Kirchner, F. (2024). Ricmonk: A three-link brachiation robot with passive grippers for energy-efficient brachiation. In *2024 IEEE International Conference on Robotics and Automation (ICRA)* (p. 8920-8926). doi: 10.1109/ICRA57147.2024.10611003

Heradio, R., de la Torre, L., & Dormido, S. (2016). Virtual and remote labs in control education: A survey. *Annual Reviews in Control*, *42*, 1-10. Retrieved from `https://www.sciencedirect.com/science/article/pii/S1367578816300281` doi: https://doi.org/10.1016/j.arcontrol.2016.08.001

Heradio, R., de la Torre, L., Galan, D., Cabrerizo, F. J., Herrera-Viedma, E., & Dormido, S. (2016). Virtual and remote labs in education: A bibliometric analysis. *Comput-*

ers & Education, *98*, 14-38. Retrieved from https://www.sciencedirect.com/science/article/pii/S0360131516300677 doi: https://doi.org/10.1016/j.compedu.2016.03.010

Javadi, M., Harnack, D., Stocco, P., Kumar, S., Vyas, S., Pizzutilo, D., & Kirchner, F. (2023). Acromonk: A minimalist underactuated brachiating robot. *IEEE Robotics and Automation Letters*, *8*(6), 3637-3644. doi: 10.1109/LRA.2023.3269296

Owens, M. T., & Tanner, K. D. (2017). Teaching as brain changing: Exploring connections between neuroscience and innovative teaching. *CBE—Life Sciences Education*, *16*(2), fe2.

Ryalat, M., Almtireen, N., Al-refai, G., Elmoaqet, H., & Rawashdeh, N. (2025). Research and education in robotics: A comprehensive review, trends, challenges, and future directions. *Journal of Sensor and Actuator Networks*, *14*(4). Retrieved from https://www.mdpi.com/2224-2708/14/4/76 doi: 10.3390/jsan14040076

Soni, R., Harnack, D., Isermann, H., Fushimi, S., Kumar, S., & Kirchner, F. (2023). End-to-end reinforcement learning for torque based variable height hopping. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 7531–7538).

van den Beemt, A., Groothuijsen, S., Özkan, L., & Hendrix, W. (2023, August). Remote labs in higher engineering education: engaging students with active learning pedagogy. *Journal of Computing in Higher Education*, *35*(2), 320–340. doi: 10.1007/s12528-022-09331-4

Wiebe, F., Babel, J., Kumar, S., Vyas, S., Harnack, D., Boukheddimi, M., ... Kirchner, F. (2022). Torque-limited simple pendulum: A toolkit for getting familiar with control algorithms in underactuated robotics. *Journal of Open Source Software*, *7*(74), 3884. Retrieved from https://doi.org/10.21105/joss.03884 doi: 10.21105/joss.03884

Wiebe, F., Kumar, S., Shala, L. J., Vyas, S., Javadi, M., & Kirchner, F. (2024). Open source dual-purpose acrobot and pendubot platform: Benchmarking control algorithms for underactuated robotics. *IEEE Robotics & Automation Magazine*, *31*(2), 113-124. doi: 10.1109/MRA.2023.3341257

Wiebe, F., Vyas, S., Maywald, L. J., Kumar, S., & Kirchner, F. (2022). Realaigym: Education and research platform for studying athletic intelligence. In *Proceedings of robotics science and systems workshop mind the gap: Opportunities and challenges in the transition between research and industry, new york.*

Zahid, M., & Pokorny, F. T. (2024). Cloudgripper: An open source cloud robotics testbed for robotic manipulation research, benchmarking and data collection at scale. In *2024 ieee international conference on robotics and automation (icra)* (pp. 12076–12082).

Workshop

Assessing student-centred learning without a written exam? A discussion on student-centred learning and continuous examination

Ana Carolina Bertassini
Technology Management and Economics (TME)
anacarol@chalmers.se

Anna Nyström Claesson
Technology Management and Economics (TME)
anna.nystrom.claesson@chalmers.se

Abstract

This contribution reports on the outcome of a workshop at KUL 2025 where a group of cross-disciplinary teachers explored continuous assessment as a means to enhance student-centred learning (SCL) in the courses. The participants were interested in student-centred learning as a means for course development and to align their courses with Chalmers' educational goals, such as constructive alignment and participatory design. During the workshop, the participants reflected on their own experiences and shared ideas on accessible learning activities. Some of the learning activities were milestone papers, authentic/performance tasks, oral explanations and mechanisms supporting engagement and self-regulation. The autonomy and independent learning in SCL have their challenges, for example, transferable skills and Generative AI. An unbalanced workload, if there is a lack of participation or institutional support, results in the students' development of autonomy and independent learning being missing, with consequences for self-regulation skills. Generative AI is another challenge and at the same time an opportunity, which calls for increased visibility and transparent tool use.

Sammandrag

Bidraget redovisar utfallet av en workshop på KUL2025 där en tvärvetenskaplig lärargrupp utforslade kontinuerlig bedömning som ett verktyg för att öka det studentcentrerade lärandet i kurserna och för kursutveckling och för att bättre möta Chalmers utbildningsmål t.ex. constructive alignment och deltagardesign. I workshopen reflekterade deltagarna själva och i grupp genom att dela med sig av idéer för bedömningsbara lärandeaktiviteter, t.ex. skriftliga avstämningar, verklighetsnära uppgifter muntliga förklaringar. Lärande aktivisterna var också mekanismer för att uppmuntra engagemang och öka självinsikten om sitt lärande. Självständigt och oberoende lärande i student centrerat lärande har sina utmaningar t.ex. överförbara färdigheter och Chat GTP. En obalanserad arbetsbelastning och om det saknas deltagande eller

stöd från lärarorganisationen kan utvecklingen av självständigt och oberoende lärande saknas kan det ha konsekvenser för utvecklingen av överförbara färdigheter. Chat GTP är en utmaning på ont och gott och uppmanar till ökad transparens och synlighet i användningen av verktyget.

1. Introduction

In the traditional classroom, the role of teachers was authoritative, and the students were passive attendants. Student-centred learning shifts the classroom power relations (Wright 2011). In an SCL classroom, the teacher's role shifts to be facilitating and the students to be responsible agents in their own learning (Wright 2011). In higher education, the interest in student-centred learning (SCL) is slowly increasing (Biggs and Tang,2011; Felder and Brent, 2016), which requires teachers to rethink roles, assessment practices, and learning design in the transition to be able to empower students to develop ownership of their learning. The structure of instruction, e.g. meaningful tasks and activities intended to create learning (Black and Wiliam 2009), requires feedback followed by purposeful assessment to contribute to complex disciplinary outcomes and to transferable skills such as metacognition, self-regulation, and collaborative competence (Boud and Falchikov, 2007; Nicol and Macfarlane-Dick, 2006).

Traditional summative assessment, such as written examinations, is useful in certain problem-solving tasks such as closed problems where the goal and method are clear (Godfrey et.al, 2014). Open problems with unclear goals and methods can be difficult to capture in written exams. The openness creates complexity, which requires complex learning such as critical thinking, the application of knowledge in real-world contexts (Biggs and Tang, 2011; Boud and Falchikov, 2007). Open and complex problems often benefit from a formative assessment where the learning outcomes are included in learning activities and continuously collected material provided by the students during a course (Black and Wiliam, 2009; Wiliam, 2011). Properly designed tasks provide feedback on the students' learning, such as low-stakes quizzes, individual and group assignments, peer reviews, reflective writing, and project-based work (Hernandez-de-Menendez et al., 2019; Black & Wiliam, 2009; Nicol & Macfarlane-Dick, 2006; Felder & Brent, 2016). The assessment provides educators with actionable information to foster self-regulation and metacognitive awareness (Dochy et al., 1999).

Useful in the design are tools such as Constructive alignment (CA) and its cousin, outcomes-based education (OBE) that specifies intended learning outcomes (ILOs) (Spady, 1994;

Harden, 2007), facilitates learner agency and participation (Lea et al. 2003; O'Neill & McMahon, 2005) and assessment strategies of achievement (Black & Wiliam, 1998; Sadler, 1989). In its formative use, continuous assessment is intended to close the gap between students' current performance and the intended outcomes through cycles of feedback and revision (Sadler, 1989; Black & Wiliam, 2009).

2. Method

The workshop at KUL 2025 was designed as a collaborative, cross-disciplinary workshop where participants could explore both learning activities and appropriate assessments. To inspire, the master's courses *Environmental Management* (VMI036) and *Stakeholder Management for sustainable development* (ENM130) were used as examples of redesigned traditional courses with summative assessment. The design of the continuous assessment in both courses was informed by research on active learning and aligned assessments (Felder & Brent, 2016; Biggs & Tang, 2011), and included elements such as quizzes, peer-to-peer feedback, group assignments and project work.

The participants were invited to outline a redesigned or new student-centred learning activity by outlining the design and continuous assessment elements, such as assignments, peer interaction, or checks for understanding. The active learning method and participatory instructional design (Könings et al., 2014) were structured using the "1-2-4-All" technique to scaffold divergent and convergent thinking. Participants first engaged in individual reflection (5-10minutes) on three prompts: (i) needs in their courses; (ii) which learning activities could/should be student-centred; (iii) what knowledge and skills were targeted in their intended learning outcomes. The individual reflections were followed by pair discussions to surface overlaps, then small group "beehives" to map alignment between outcomes, activities, and assessment tools. Finally, a plenary synthesis to share designs, constraints, and next steps.

Guiding questions were: RQ1, how can continuous assessment be designed to elicit valid evidence of complex outcomes in SCL courses? RQ2, in what ways does continuous assessment support students' engagement, self-regulation, and learning? RQ3, what affordances and constraints do educators identify when adopting continuous assessment?

The data sources comprised (a) facilitator field notes from pair, beehive, and plenary discussions, and (b) group summary sheets. The data was analysed with qualitative analysis and organised as summaries for each research question. For RQ1, we recorded whether activities showed clear outcome–task–evidence alignment and any authenticity features. How the participant developed and addressed design features eliciting valid evidence of a complex outcome. RQ2 noted mechanisms that supported engagement and self-regulation (e.g., goal/criteria clarity, feedback timing, self/peer assessment opportunities and revision). In RQ3, the perceived advice and limitations by the instructors are examined. The use of tallied simple frequencies (e.g., number of activities showing alignment) and selected illustrative excerpts was used in the evaluation of the data for accuracy and completeness.

3. Results

The participants recognised the potential for continuous assessment both as a learning process and as an assessment mechanism. Activities such as quizzes, discussions, reflective writing, group projects, and peer review were already used in various forms by the participants, but were rarely used as assessment instruments. To better support critical thinking, collaboration, and knowledge application, the participants considered formalising and aligning the activities with course learning outcomes. During the workshop, the participants identified existing practices, unmet needs, and alignment opportunities between learning outcomes and assessment strategies. Key findings by course are summarised below.

Course 1 – Integration of Active and Reflective Practices

Participants teaching Course 1 wanted to refine how group work translated into deeper learning, including peer accountability and with a stronger connection to the project's course objectives. Student-centred strategies were described at individual, group, and in-class levels. In-class sessions included interactive discussions, mini-exercises, and so-called "puzzlers", short problem-solving prompts designed to stimulate critical thinking and immediate engagement. The students engaged individually with preparatory readings and completed exercises, occasionally using a flipped-classroom format. In supporting self-reflective activities, the participants used low-stakes quizzes. Collaborative activities, e.g. research, reporting, and presentations, were deliberate efforts to increase student ownership of real-world applications.

Course 2 – Emphasis on Engagement and Thematic Understanding

Participants teaching Course 2 expressed a need for enhanced discussion, deeper student engagement, and more continuous assessment activities. Existing practices partially addressed these goals, but additional structures were considered necessary to help students process key readings and consolidate their understanding of core themes. Participants proposed recurring exercises and scaffolding of group tasks to strengthen students' responsibility for their learning.

Suggested adjustments included weekly discussion prompts, peer-explained summaries, and group assignments with increasing complexity, subject to continuous assessment.

Course 3 – Mixed Modalities and Foundational Learning

The participants identified the potential to increase students' ownership of learning through additional active-learning elements. Proposed activities included short writing tasks, peer assessments, and reflective components following quizzes, with digital tools to personalise feedback and enhance formative learning opportunities. The course combined lectures with in-class, basic, closed problem-solving exercises in pairs. Mini-projects in a seminar setting offered application of concepts and peer-learning opportunities. Online quizzes were used to provide feedback and monitor understanding.

Course 4 – Focus on Independent Work and Final Integration

The assessment in course 4 relied more on summative than formative approaches, centred on individual assignments and a final project. The new design included extending existing project work into a continuous assessment process with embedded checkpoints (e.g., milestone submissions, peer reviews, and progress reflections). Learning logs, oral presentations, and peer commentary during project development could be used in the formative feedback to reduce the focus of the final report product.

4. Discussion

A common thread across the discussions was the importance of feedback as a key feature of effective continuous assessment, both from peers and instructors. Consistent with prior work, timely and specific formative feedback helps students regulate their learning and deepen their understanding of course content (Nicol and Macfarlane-Dick 2006). The workshop revealed a keen interest among educators in student-centred learning and continuous assessment as means to support engagement, thinking skills, knowledge development, collaboration, and problem-solving. Participants identified instructional strategies that aligned student-centred learning activities with collaborative research and reflection.

Participants (RQ1) consistently pointed to designs that made the outcome–task–evidence relationship explicit and auditable. In line with the student-centred framing, participants also emphasised the development of transferable skills (e.g., metacognitive monitoring, self-regulation, collaboration, and accountable participation), which are frequently cited outcomes of feedback-rich designs (Nicol and Macfarlane-Dick 2006). However, there was a shared view that more explicit integration of these activities within assessment frameworks was needed to ensure constructive alignment and to strengthen the evidential basis for complex outcomes (Biggs & Tang, 2011; Wiggins, 1998).

Concerning RQ2, these accounts indicated that continuous assessment supported students' engagement and learning by cultivating skills, recurring feedback opportunities, reflection, and revision. Activities such as in-class discussions, peer-reviewed assignments, and progress tracking in project work reflect recommendations to embed assessment within the learning process (Felder and Brent, 2016). These approaches encourage students to take ownership of learning, build metacognitive awareness, and develop collaborative skills, which were outcomes increasingly emphasised in contemporary higher education (Boud and Falchikov, 2007). Taken together, these patterns offer a direct answer to RQ2, that continuous assessment supports engagement and self-regulation by orchestrating feedback–revision cycles that make learning processes visible and improvable.

Yet, practical challenges were reported, for example, time constraints, inconsistent student participation, and the need for institutional support when implementing alternative assessment practices, anticipating RQ3 on advice and constraints. Consistent with the literature, participants viewed continuous assessment not merely as a grading mechanism, but as a pedagogical approach that supported learning through regular feedback, iterative development, and formative check-ins (Black & Wiliam, 2009; Nicol & Macfarlane-Dick, 2006).

Course-specific reflections reinforced the interpretation and illustrated answers to RQ1 and RQ2. Reflections from Course 1 underscored the value of active learning through in-class problem-solving and group research. To translate learning activities into meaningful assessment opportunities in line with constructive alignment (Biggs and Tang's 2011), ensuring a clear match between intended learning outcomes, tasks, and assessment (evidence) can be a challenge. In course 2, the need for stronger support mechanisms to help students engage with complex readings and key themes suggests that recurring exercises and discussion-based learning can guide learning behaviour over time (Gibbs and Simpson 2004). For Course 3, incorporating problem seminars and online quizzes was perceived as a pathway toward more active and formative learning; proposed peer-based and reflective elements align with designs shown to foster student agency (Dochy et al., 1999). In Course 4, assessment was primarily summative, centred on individual assignments and a final project. There was a need to break assessment into smaller, scaffolded components resonating with the formative principle of helping students closing the gap between current performance and intended learning outcomes through cycles of feedback and revision (Sadler, 1989).

In RQ3, participants also raised generative AI as both an opportunity and a challenge for continuous assessment. Because short written products, summaries, and early drafts can be AI-assisted, concerns focused on the visibility of authorship, the displacement of learners' cognitive effort, and equity in tool access and disclosure, issues widely noted in higher-education discussions (Zawacki-Richter et al., 2019; Kasneci et al., 2023). The trajectory highlighted in the workshop, toward process-anchored assessment (iterative checkpoints, feedback-to-revision cycles, and occasional in-class/explain-your-work moments), was well aligned with institutional guidance prioritising process over product. The pragmatic response to make the learning processes visible, was to reduce the reliance on single artefacts vulnerable to AI support, by sustaining validity for complex outcomes (Biggs and Tang, 2011). At the same time, AI foregrounds design tensions for continuous assessment (e.g., workload, monitoring outside class, transparency of tool use), clarifying both the constraints to be managed and the affordances to be leveraged (RQ3).

Adapted to the realities of AI-enabled study practices (Boud & Falchikov, 2007; Nicol & Macfarlane-Dick, 2006; Kasneci et al., 2023), suggests that continuous assessment, when deliberately aligned with outcomes and supported by feedback-rich routines, can create complex learning and transferable skills (metacognition, self-regulation, collaboration).

5. Conclusion

In this contribution, continuous assessment has been explored as a meaningful alternative to traditional written examinations in student-centred learning environments. The workshop's multidisciplinary group of educators converged on aligning assessment strategies with active and reflective learning processes. Substituting or supplementing summative exams with continuous assessment makes outcomes more visible, achievable (a sequence of alignment, feedback-oriented tasks), deepens learning, and strengthens student autonomy and self-regulation. In particular, open-ended, real and authentic problems can constructively be aligned to intended outcomes (Sadler, 1989; Nicol & Macfarlane-Dick, 2006; Biggs & Tang, 2011).

Active learning and continuous examination carry challenges, including workload concerns, institutional constraints, and varying levels of student readiness for self-directed learning. In an AI-enabled learning environment, many tasks in the continuous assessment (short writes, summaries, early drafts) can be AI-assisted. To sustain validity, authenticity and transparency in the assessment, processes need to be accompanied with documented revision histories, brief explain-your-work moments, and transparent AI-use expectations. As continuous assessment becomes more embedded in educational practice, faculty development efforts should focus on strategies for sustainable implementation, effective feedback design, and learning outcome alignment in the disciplinary contexts. Assessment is not separated from learning; it is integrated. Through intentional design and reflective practice, continuous assessment can transform both teaching and learning, moving higher education closer to the goals of inclusivity, relevance, and lifelong learning.

References

Biggs, J., & Tang, C. (2011). Teaching for Quality Learning at University (4th ed.). OpenUniversity Press.

Biggs, J. & Tang, C. (2020), Constructive Alignment: An Outcomes-Based Approach to Teaching Anatomy, (2nd ed), https://doi.org/10.1007/978-3-030-43283-6 © Springer Nature Switzerland AG 202

Black, P., Wiliam, D. Developing the theory of formative assessment. Educ Asse Eval Acc 21, 5–31 (2009). https://doi.org/10.1007/s11092-008-9068-5

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. Assessment & Evaluation in Higher Education, 31(4), 399–413. https://doi.org/10.1080/02602930600679050

Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. Studies in Higher Education, 24(3), 331–350. https://doi.org/10.1080/03075079912331379935

Felder, R. M., & Brent, R. (2016). Teaching and Learning STEM: A Practical Guide. Jossey-Bass.

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students'learning. Learning and Teaching in Higher Education, 1(1), 3–31

Hernández-de-Mendéndez M., Vallejo Guevara A., Tudón Martínez J. C., Hernández Alcántara D., Morales- Mendéndez R, (2019) Active learning in engineering education. A review of fundamentals, best practice and experience. International journal of Interactive design and Manufactoring (2019),13, 909-922

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and individual differences, 103, 102274.

Könings, K.D., Seidel, T. & van Merriënboer, J.J.G. Participatory design of learning environments: integrating perspectives of students, teachers, and designers. Instr Sci 42, 1–9 (2014). https://doi.org/10.1007/s11251-013-9305-2

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. Studies in Higher Education, 31(2), 199–218. https://doi.org/10.1080/03075070600572090

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18(2), 119–144

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? International journal of educational technology in higher education, 16(1), 1-27.

Wright G.B. (2011), Student-centred learning in Higher education., International Journal of Teaching and Learning in Higher Education, 23(3) 1812-9129

KUL
2025

KUL 2025 – Chalmers Conference
on Teaching and Learning

The conference is aimed at Chalmers' employees and students and is about all aspects of higher education; teaching, learning, supervision, educational leadership, and student relations.