ASSESSING STUDENTS OR THEIR AI TEAMMATES? CASE STUDIES IN THE GAI ERA

Jens Bennedsen

Department of Electrical and Computer Engineering, Aarhus University

Joo Ghee Lim

School of Electrical & Electronic Engineering, Singapore Polytechnic

Mark Nivan Singh

Department of Educational Development, Singapore Polytechnic

Mika Suutari

General Administration, Turku University of Applied Sciences

Reidar Lyng

Center for Science & Engineering Education Development,
Norwegian University of Science & Technology

Siegfried Rouvrais

Lab-STICC CNRS 6285, IMT Atlantique

ABSTRACT

The widespread adoption of Generative AI (GAI) since 2022 has fundamentally challenged traditional assessment methods in engineering education, exposing the weakness in assuming submitted artefacts reliably represent student thinking. This article's purpose is to provide concrete examples of how GAI has changed formative and summative assessment in engineering courses. Through comparative case studies across five CDIO institutions in Norway, Denmark, Finland, France, and Singapore, we examine GAI implementations involving over 400 students. All cases are characterized using the Traffic Light Model by Arene (2024), which categorizes AI use as required (blue), prohibited (red), reported (yellow), or allowed (green). Our findings reveal that effective GAI integration requires thoughtful assessment redesign—structured implementations with clear pedagogical frameworks demonstrated positive outcomes, whilst less structured approaches can lead to student over-reliance and reduced learning. The evidence indicates thoughtful assessment redesign grounded in constructive alignment is required, where AI use decisions are tied explicitly to learning outcomes. Effective integration demands shifting from product-based to process-based assessment methods, emphasizing evaluation of student reasoning, judgement, and capacity to work responsibly with AI.

KEYWORDS

Generative AI, assessment, learning paths, CDIO Standards 2, 3, 11.

INTRODUCTION

Since the release of ChatGPT in November 2022, Generative AI (GAI) has been on everybody's lips. Especially in teaching and learning, it has been clear that the traditional assessment formats are deeply challenged by GAI. This is particularly so in Higher Education (HE), where essays, technical and managerial reports, programming problems, math problems, etc. can all be easily solved and answered by one of the freely available GAI tools. Recently, we have seen a bigger integration of the tools in many software products, e.g. an AI answer when using Google Search, correction of grammatical errors and suggestions for improvement of a text in your text editor.

Many of the GAI tools are free of charge and very easy to use; thus, many students found these tools attractive and started using them. Sakirin and Said (2023) "*found that 70% of users chose ChatGPT-powered conversational interfaces over traditional techniques, citing convenience, efficiency, and personalization*." (p.22). Although many institutions banned these tools in the beginning, many have since shifted towards a more positive view of seeing GAI as an effective learning tool (McDonald, Johri, Ali, & Hingle Collier, 2025). In 2026, more and more Universities start to have their own sovereign RAG LLM servers.

Especially within assessment, the GAI tools have challenged our conventional ways. As Niels Lauritzen (2025) notes in his debate post "*With a simple prompt, uploading my notes and exam assignments to Google AI Studio, I saw for the first time generative AI produce a correct answer with completely correct references to the syllabus*.". Jens Bennedsen (2025), in his reply to Niels Lauritzen's post, wrote "*We evaluate the students' process (how they solve a task) by evaluating their product. The problem is that with GAI, we now have a digital assistant that can produce a product of a quality that will typically result in the highest grade. In other words, we cannot assume a correlation between the students' process and the product they create. What is the solution? We need to develop types of examinations that allow us to evaluate the students' processes directly. This requires that we can observe the process or at least get the student to discuss it based on questions such as: What did you do to make the product? What alternatives are there? What considerations led you to choose this product? Why does it make sense to use this theory?*"

The Traffic Light Model published by the Rectors' Conference of Universities of Applied Sciences in Finland (ARENE, 2024) categorizes AI use as required (blue), prohibited (red), reported (yellow), or allowed (green) (Figure 1). Students at universities of applied sciences are recommended to use AI in learning and studying. The model is used in Finnish Universities of Applied Sciences to clearly communicate when and how the use of AI tools applies to course assignments (Kontio, 2025).

Figure 1: The labels of the Traffic Light Model.

Based on the Traffic Light model, this paper gives examples from teachers at five CDIO institutions on how they allowed and used GAI for formative and summative assessment in 2025. A short pedagogical discussion on the assessment of product or process in the GAI new era is given, then impacts on educational frameworks for quality enhancement (as accreditations and CDIO) are questioned.

RELATED WORK

Growing student involvement in assessment and feedback with GAI

The growing use of GAI by students is changing how assessment and feedback take place. For teachers, this phase is often stressful and time-consuming. Could teachers make their workload lighter and boost student learning by guiding students to evaluate their own outcomes and engage in peer review with GAI?

Research by Zhan et al. (2025) suggests that GAI has the potential to enhance student feedback engagement with its rapid personalized feedback, ongoing interaction and practice. Nonetheless, significant challenges and ethical concerns also arise. Without proper training, students might misuse GAI, which could inhibit their engagement and learning. To support students, teachers should assist students in defining their individual feedback goals before seeking feedback and create ample opportunities to interact with GAI tools.

While students generally consider teacher feedback to be more helpful and trustworthy, GAI feedback stands out for being much easier to access, more understandable, less negative, and better at aiding sense-making and understanding. The findings suggest that these feedback modes are not an either/or; students can and do engage with both, because they serve different functions (Henderson, et al., 2025).

FIVE CASE STUDIES IN ENGINEERING EDUCATION

The following case studies shed light on different ways to tackle the “problem” that GAI has made clear when assessing students – are we assessing the students or the GAI?

Developing Critical Thinking Practices Interwoven with Blue GAI usage at NTNU

In a study published by Styve et al. (2024), a first-year bachelor's 7.5 ECTS engineering course designed to foster critical thinking skills was adjusted to integrate the use of GenAI. Present day software development is largely characterized by assembling premade blocks of code, rather than creating and coding algorithms. If students perceive the task to be the delivery of a well-functioning code, there is a risk that coding blocks found on internet forums are uncritically assumed to perform what they claim to do. The recent establishment of generative AI tools exacerbates this challenge.

The course design - formative assessment

For the study, one of four mandatory assignments in the course was redesigned and titled “AI with Critical Thinking”. (In addition to the four assignments, students had to finish and deliver a portfolio project.)

The students were instructed to solve a programming assignment by using GitHub Copilot to generate code, and then analyses the code suggested by Copilot with respect to the quality and robustness criteria taught in the course. They were then instructed to ask ChatGPT to analyses the code generated by Copilot, compare the output with their own analysis, and write down their observations and reflections on the process so far. Based on their own analysis and that of ChatGPT, they were then instructed to refactor the generated code to the required level of quality and robustness. For the final part, they were also required to describe and explain their changes to the code and answer questions to reflect on the overall experience.

Results

Before and after this assignment, changes in the students' Critical Thinking skills were evaluated using a Critical Thinking Survey Questionnaire. The sample size was 90 students, with 55 that filled out the pre/post survey. In addition, 88 of 90 students submitted their assignments with reflective questions included. Overall, the results indicate that the students' awareness of the limitations of AI Tools, the importance of code quality and robustness, and of critical thinking practices were raised by the redesign.

The redesigned assignment did raise awareness among the students, as seen from the survey and the thematic analysis of reflective questions. The authors report that the positive effects on critical thinking from the pre- and post-surveys were positive after a single iteration. The most notable change in the survey was for the claim "*I use a piece of code in my program even if I do not fully understand it*" which declined from 'seldom' closer to 'never'. The reflective questions clearly indicate an increased awareness among the students with respect to limitations in the trustworthiness of AI. For further in-depth analysis the reader is referred to the original paper.

Accessing the programming process instead of the program using blue-green GAI at Aarhus University

One of the areas GAI is believed to deliver a significant reduction in the man-hours needed is programming. The LLMs are trained on a huge amount of source code etc., making them very good at producing code for the kind of problem very often used in introductory programming courses.

The context

An introductory programming course (5 ECTS) at a professional bachelor's programme (3,5 years). The class size is around 175. The students are divided into eight groups of 20-25 students. There is a lecture (2*45 minutes, all students in an auditorium) and problem-solving class (2*45 min in their group) each week. There is a final exam at the end of the course with a pass/fail grade.

Previously, the final exam was a "traditional" written exam containing programming problems where the students should implement programs solving the problems. The grading of their programming competencies was based on the product, i.e. the source code. Any of the GAI tools can create flawless implementations of the programming problems, thus making it problematic to use their product as a proxy for their competencies.

The assessment

The students use GAI to create a solution (program) to the problems. Instead of using the source code as the basis for the assessment, the student must explain the program and how it solves the problem, i.e. it has changed from a written to an oral exam.

The students get 4-5 programming problems to solve. The time is the same as the previous written exam (two hours). Ten to twelve students sit in a large room at the same time. The examiner and co-examiner walk around in the room and ask the students about their solution and make small modifications (without the use of GAI) to their code. In this way, 10 students can be examined in two hours, making the form as efficient as the traditional written exam (previously, it took approximately 10 minutes to grade a student's solution).

Student feedback

After the first time we held this type of exam, we evaluated it with the students by sending them a questionnaire. It was a rather small cohort (approximately 70 students), and the response rate was 47%. In general, the students found that the form evaluated the course's learning outcomes; it was less stressful than a normal exam and resembled the way they envision programming will be in the future (you must understand and modify code made by GAI). On the negative side, the students would have liked more time to show their solutions to the examiner.

Teacher feedback

It was an effective exam, but quite stressful to keep track of 10-12 students at the same time. It took some time for the students to understand what we wanted when asking them about their solution (not "read C++ code aloud but explain the code using the correct terminology). We have therefore developed a more precise rubric to assess each student. To enhance their explanation skills (evaluative judgement of code produced by GAI), the problem-solving classes focus more on students explaining code to each other.

ICT tools in green-yellow-red GAI at Turku UAS

This case examines the implementation of GAI in a first-year engineering course, ICT Tools (2 ECTS), at Turku University of Applied Sciences. The approach reflects a shift toward a new learning culture emphasizing hybrid intelligence, shared responsibility in assessment, and the importance of AI literacy.

Pedagogical framework

Traditionally, assessment relies on teacher evaluation and feedback after submission. From a learning perspective, it is essential for students to analyze the feedback they receive and learn from it. Unfortunately, this process is often minimal, and students lack motivation after submission. In the model used in this course, AI acts as a mentor and sparring partner, enabling iterative improvement before final submission. Students engage in AI-assisted self-assessment based on the assessment criteria provided by the teacher, and maintain learning diaries, making the learning process transparent and enhancing learning and metacognitive skills.

At the beginning of the course, the teacher presents a slide with a title "AI as a mentor and sparring partner" among other topics, including instructions on peer review: "You can also use AI in peer review of your own performance. If necessary, give AI the evaluation criteria according to which your output should be evaluated."

Instructing the responsible AI use

At Turku UAS, the implementation of the traffic light model is part of the overall process of promoting and activating the use of AI in teaching and learning and strengthening AI literacy (Kontio, 2025). The Traffic Light Model is used in the instructions of each learning task to promote equity, AI literacy, and reduce fraud-related concerns. For example, independent weekly tasks in this course allow unrestricted AI use (green), while specific assignments require reporting (yellow), including prompts, AI outputs, and self-evaluation. In the final exam, the use of AI is prohibited (red).

A student example: good and bad practices

As an independent PowerPoint assignment for the ICT Tools course, the student can submit a presentation on a professional topic of interest. The assignment has a yellow traffic light. Before submitting the assignment, the student must check that the presentation contains all the elements required in the assignment and prepare a self-assessment using a ready-made assessment framework. The assessment is affected by whether the presentation contains the required elements. The task focuses on creating functional presentations that include the required elements. The process can be AI-assisted.

A student tested Microsoft 365 Copilot to generate an eight-slide presentation aligned with course criteria. While the output was visually appealing, it lacked required elements (hyperlink, table, chart, video, and graphic objects). The student gave Copilot an assessment framework used by the teacher, and as a result, Copilot told the student about the shortcomings and gave development suggestions. However, the student submitted the AI-generated version with self-evaluation but without revisions, resulting in a lower grade. In an ideal situation, the student would have made the changes after receiving the AI peer review, for example, with the help of AI sparring, learning in the process, and finally submitting the improved presentation, which would have led to a higher grade. This would also have reduced the teacher's workload in the final assessment and final feedback.

Project management planning using blue GAI at IMT Atlantique

The MGP210 “Project Management Methods” course spans 16h30 over two months. Offered since 2007, it is mandatory for the 60 engineering students enrolled in the IT apprenticeship track during their first Master’s year. The cohort is divided into three classes of 20 students, organized in teams.

The course is structured on a project-based learning model (11h15), in which students first delineate a project idea and match it with a real funding call (e.g. regional, Horizon Europe). Then they elaborate a project management plan (no implementation).

Seven intended learning outcomes (LO) are used for assessment: (LO1) identify and formalize project objectives, strategic issues, main activities and deliverables, (LO2) define a work breakdown structure with task assignment, (LO3) estimate effort & task durations in a schedule sequencing activities, (LO4) identify, qualify and quantify project risks, with mitigations, (LO5) define SMART KPIs, (LO6) integrate environmental, social-societal dimensions as financial & legal concerns, and (LO7) structure and align project concerns in a logical framework. The grading model is competency-based, students are assessed for each LO with a rubric on a scale: not observed, not acquired, partially acquired, acquired, and beyond expectations. For crediting, three LOs must be acquired or higher, and none at not acquired or below.

In the 2024 session, 80% of the grading was based on student’s project management plan reports (approx. 60 pages) and 20% on a final oral presentation. Due to the increasing use of GAI by students in their reports, as an experiment, it was decided in Fall 2025 to let students an ‘open bar’ on the usage of GAI (Traffic Light Model “use of AI required”). Students elaborate their report with GAI somehow as a knowledge team member. The grading moved to 60% for oral and 40% on report. Still in 2025, the institution did not yet have GAI regulations formalized, and no GAI introductory course is in place.

In 2025, the assessment was threefold. First, an intermediate oral presentation (flipped-classroom model) of teams on their understandings of methods and tools on LO2-6 (1h15 for

5 teams of 4, one team per LO2-6 roles) is fully formative on lower Bloom levels. Slides are mostly generated with an AI tool by students. Secondly, a final oral presentation with Q&A (1h15 for 4 teams of 5) is summative, where both the teacher (50%) and students assess their peers (50%) on higher Bloom levels. This presentation allows for a more objective verification of skills. Via an online tool, each team is assessed on the seven LOs by another class of 20 students attending the presentations. Students also provide textual comments on each. Finally, the reports are assessed by the teacher with the help of GAI (Mistral 7B-instruct on a personal laptop), for 40% of the final mark. Each team report follows the same structure thanks to a LaTeX template of the project management plan, with a section for each LO1-6. LO7 is transversal. Selected document sections are ingested as a knowledge base (PM-BoK, Open PM² developed by the EU commission, and some dedicated ISO). Educationally, structured frameworks for LO1-LO7 prompts are used to guide assessment with the rubric criteria, 7W2H and PREPARED4ED (Verkuilena, 2025).

This sandboxed course clearly has no quality evidence of the use of GAI, it needs some pedagogical realignment in the Biggs sense. An evaluation questionnaire was filled by students at the end of the course. Quantitatively, their appreciation of full freedom to use GAI was rated above 7/10, but qualitatively, some students gave critical comments. In 2025, teams used mainly ChatGPT and Claude, which rely on a massive amount of data, and tend to limit students' reasoning via copy-and-paste answers. For the Fall 2026, a project management planning chatbot would pedagogically help to ensure coherence between guided responses and each LO. As a consequence, a chatbot is to be designed and a dedicated RAG with chunks from the knowledge base is to be envisioned.

Formative flipped learning assessment using yellow GAI at Singapore Polytechnic

Singapore Polytechnic has reduced reliance on high-stakes examinations by strengthening AI-enabled formative assessment as a core mechanism for preserving student cognition while maintaining assessment robustness. The institution implemented an AI-supported formative assessment model embedded within synchronous flipped-learning lessons. The approach was piloted in 2025 with 59 first-year diploma students across three classes, taught by three experienced lecturers, providing a realistic instructional context.

During each lesson, students completed an initial in-class quiz generated from approved learning materials. Conducted under supervised conditions, (1) first, this quiz functioned as a low-stakes, high-signal formative assessment, capturing authentic evidence of conceptual understanding. GAI analyzed student responses in real time, identifying misconceptions and patterns of difficulty at both individual and class levels. Lecturers then (2) adjusted instruction through just-in-time teaching, reallocating lesson time to address diagnosed gaps rather than pre-planned content coverage. A second in-class quiz followed, allowing both lecturers and students to observe immediate learning progression. Analysis of combined quiz results showed clear improvement for multiple students between the first and second quizzes, indicating measurable cognitive gains within a single lesson cycle. At the end of the lesson (3), GAI generated personalized learning reports for students, including targeted notes and follow-up tasks derived strictly from demonstrated performance. These reports were designed to support reflection and improvement, not to replace student thinking.

In this design, GAI functions as a cognitive amplifier rather than a cognitive substitute. This model preserves cognition in three ways. First, supervised, real-time sensemaking requires students to interpret questions and apply disciplinary principles without extended opportunities for AI outsourcing. Second, metacognitive regulation is reinforced through personalized feedback grounded in students' own errors and misconceptions; 72% of students used the personalized notes, and 97% of these reported positive learning impact. Third, differentiated cognitive demand ensures foundational mastery for weaker students while extending higher-order application for stronger students, avoiding the cognitive flattening common in uniform assessments.

From an assessment quality perspective, the approach strengthens robustness across multiple dimensions. Task design aligned closely with learning outcomes enhances validity. Repeated formative checkpoints and aggregated performance patterns improve reliability. Supervised administration and process-level evidence (attempts, progression, misconceptions) strengthen academic integrity. Diagnostic reports and performance trends provide defensible evidence to support summative decisions where required.

Student perception data reinforces acceptability and feasibility. With an 80% survey response rate, over 80% of students rated the approach as engaging and effective, and 94% valued the structured quiz–lesson–quiz–feedback sequence, suggesting that cognition-preserving assessment design was well received rather than resisted.

This use case demonstrates that the response to GAI-disrupted examinations lies not in tighter surveillance or blanket AI prohibitions, but in intentional assessment redesign. By embedding AI-enabled formative assessment at the heart of teaching and learning, SP preserves student cognition, enhances instructional responsiveness, and maintains assessment robustness in an AI-rich educational environment.

DISCUSSION

This paper is twofold: experiences of GAI tools usage by engineering students and of GAI usage by faculty for their student formative and summative assessments. GAI disrupts a core assessment assumption: that the submitted artefact is a reliable proxy for the learner's thinking. In GAI-rich contexts, product quality alone becomes weak evidence of competence, shifting the problem from integrity to validity, specifically what claims can defensibly be made from the evidence collected (Zhao et al., 2024). Across the five cases, the practical implication for CDIO is consistent: assessment needs to generate process evidence, not only a final product.

What engineering students are doing with GAI, learning outcomes considerations

Graduate outcomes are disciplinary knowledge oriented, but also transversal for engineering professional competences. GAI as a teammate displays transversal skills: capabilities to understand, apply, analyze, evaluate, and create. It can offer instant results (e.g. text, source code, slides), but also guidance in student learning paths and journeys, as a teammate chatbot or a process-oriented formative sparring partner. As Stokkink (2025) concludes in the scope of GAI, "*proper adjustment of assessment methods to align with learning objectives is essential. [...] formative and summative assessment should be aligned*". All the case studies show that AI literacy and critical thinking are key competencies in the modern era of AI, and cognitive processes should not be outsourced to GAI. Hybrid intelligence (human-GAI collaboration) seems to optimise learning outcomes. This reframes competence as engineering with and around AI, including error detection, verification, and responsible disclosure (Zhan et al., 2025).

GAI with an assessment role: the process rather than the result

GAI has fueled the discussion of who the actual "person" is that produces the product we assess. But, did we previously know if it really was the student who created the product? One could argue that GAI has just changed “ghost-writing” (others creating the assessed product) from something only accessible to students with financial capabilities to all now.

Rethinking the roles of the student and the teacher has great potential. AI can enhance formative feedback, improve learning and grades, and reduce teacher workload, but risks require careful management and guidance. Clear guidelines on GAI use in learning tasks are essential to ensure equal and fair assessment and reduce fraud-related concerns. The Traffic Light Model (ARENE, 2024) on an organisational level is an example of a useful tool that supports shifting assessment toward critical thinking and judgment rather than end products alone (Kizilcec et al., 2024).

Feedback from students at IMT Atlantique case study in 2025 highlights several concerns (translated verbatim): “Abolish the 100% AI in favor of a better balance of personal reflection / AI writing”; “AI does the job for us too well […] we did not retain much of the concepts because they relied too much on AI, while the concepts covered are crucial and very relevant in business”, “the solution that we ended up finding, the AI didn't have it for us and couldn't have offered it to me”. This is illustrated also by student feedback: “starting from 0 with the AI, I could only accept what it said without critical thinking since I had no hindsight on what it said or knowledge. We should learn the fundamentals and then use AI as a tool rather than making it a substitute”. Another replied: “perhaps you could give us examples of relevant questions to ask so that we can better understand how to engage with the tool and what can actually be expected from it”.

Navigating the assessment, traffic lights of the 5 case studies

In the case of AU, the focus is on programming as a verb rather than a noun, i.e., the programming process rather than the program. The primary focus is students' demonstrated understanding of program logic, their ability to reason about behaviour, and their capacity to make targeted modifications without relying on GAI.

In the case of SP, the focus on formative assessment allows students to use GAI that provides feedback and pathways to enhance their learning under the supervision of the lecturers. While not affecting grades directly, open and green approaches remain credible when paired with designs that surface process, such as oral defence, staged deliverables, and critique-and-improve tasks.

In the case of NTNU, the students are required both to use GAI (blue light), and to comment on the use and how the output from the tools relates to their assignment (orange light), in line with recommendations by Haugan et al. (2017) to redesign mandatory coursework assignments as formative assessment exercises. One aspect of further work that may need to be addressed is feedback engagement. Feedback alone does not have intended effects without student engagement (Zhan et al., 2025). Another way of describing the same challenge is the phenomenon dubbed the “reflective zombie”, as outlined by de la Croix and Veen (2018).

DISCUSSION

AI is a disruption happening in a very rapid pace. In 2026, some institutions have GAI regulations, while others do not. Some are already engaged in their conformance with, e.g., the EU AI Act (data confidentiality, sovereignty, etc.), while others do not. Student and faculty AI literacies are at their early stages. GAI is a fast-moving topic where students in many cases are as competent as their teachers. QA agencies and accreditation bodies (e.g., US ABET, French CTI, EA, AACSB) for the higher education sector will take time to revise their reference criteria on GAI. Formal mandates might soon emerge. As of today, and to the best of our knowledge, in 2026, most international accreditation bodies for engineering education only prompt institutions to regulate their students and faculty's GAI use via internal policies.

Integrating GAI in educational programs is no small feat. It will take time to reach a coherent way of doing that free of inconsistencies. For example, within the non-normative CDIO framework several of the standards are affected by the impact of GAI use. The *CDIO Standard #2* needs to cover GAI use by students in their learning trajectory, both with respect to mentoring, and as an engineering and reporting tool. The *CDIO Standard #11*, covering the assessment of learning, needs to be explicit about assessment of processes rather than end products alone. Effective and reliable student assessments should use a variety of methods, constructively aligned with each type of outcome (Stokkink, 2025; Corbin et al., 2025). Furthermore, the *CDIO Standard #10* on enhancement of faculty competence will need to be revisited. Educators and teachers, disciplinary and pedagogical experts, need not be LLM specialists but they do need to figure out how GAI can be used to enhance learning. We suggest that we – the CDIO community – start revising our standards.

CONCLUSION

Generative AI has exposed a fundamental weakness in conventional assessment: the assumption that the quality of a submitted artefact can be treated as reliable evidence of student learning. Across the five cases presented, a common implication emerges—in GAI-rich environments, assessment must generate credible evidence of student thinking, including reasoning, judgment, and the capacity to work responsibly with AI.

The cases show there is no universal response. What is required is thoughtful assessment redesign grounded in constructive alignment, where decisions about when AI should be prohibited, allowed, reported, or required are tied explicitly to learning outcomes. Whether through oral examinations, reflective critique, or supervised tasks, the direction is clear: engineering education must increasingly assess how students think and learn, not only what they produce.

From a CDIO perspective, this extends beyond individual courses to programmatic alignment of GAI-related learning outcomes, faculty capability, and assessment practices across curricula. The central question is no longer whether engineering education should respond to GAI, but how it can do so while preserving cognition, strengthening professional judgement, and sustaining assessment validity in an AI-rich era.

ACKNOWLEDGEMENT

The authors received no financial support for this work.

REFERENCES

ARENE (2024). Arene’s recommendations on the use of artificial intelligence for universities of applied sciences. From https://arene.fi/julkaisut/arenen-suositukset-tekoalyn-hyodyntamisesta-ammattikorkeakouluille/

Bennedsen, J. (2025, September 29). Opinion: GAI is Forcing us to Develop New Types of Examination. From https://omnibus.au.dk/en/archive/show/artikel/debat-gai-tvinger-os-til-at-udvikle-nye-eksamensformer

Corbin, T., Bearman, M., Boud, D., & Dawson, P. (2025). The wicked problem of AI and assessment. *Assessment & Evaluation in Higher Education*, 1–17. https://doi.org/10.1080/02602938.2025.2553340

de la Croix, A., & Veen, M. (2018). The reflective zombie: Problematizing the conceptual framework of reflection in medical education. Perspectives on Medical Education, 394 - 400.

Haugan, J., Lysebo, M., & Lauvas, P. (2017). Mandatory coursework assignments can be, and should be, eliminated! European Journal of Engineering Education, 1408–1421.

Henderson, M., Bearman, M., Chung, J., Fawns, T., Buckingham Shum, S., Matthews, K. E., & de Mello Heredia, J. (2025). Comparing Generative AI and teacher feedback: student perceptions of usefulness and trustworthiness. Assessment & Evaluation in Higher Education, 1-16.

Kizilcec, R. F., Huber, E., Papanastasiou, E. C., Cram, A., Makridis, C. A., Smolansky, A. Zeivots,S.& Raduescu, C., (2024) Perceived impact of generative AI on assessments: Comparing educator and student perspectives in Australia, Cyprus, and the United States. *Computers and Education: Artificial Intelligence, vol 7*

Lauritsen, N. (2025, September 23). Opinion: AU's GAI Recommendations are already a Year Old – but have already been overridden by technology. From https://omnibus.au.dk/en/archive/show/artikel/debat-aus-gai-anbefalinger-er-et-aar-gamle-men-er-allerede-overhalet-af-teknologien

Kontio, J. (2025). Implementing AI traffic lights for teaching and learning – case Turku UAS. ISATE2025. From https://isate2025.toyota-ct.ac.jp/wp_system/wp-content/uploads/2025/09/05-46-ST3-4FI_01.pdf

McDonald, V., Johri, A., Ali, A., & Hingle Collier, A. (2025). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. Computers in Human Behavior: Artificial Humans, 3. doi:https://doi.org/10.1016/j.chbah.2025.100121

Perkins, M., Roe, J., Furze, L., & MacVaugh, J. (2025). Reimagining the Artificial Intelligence Assessment Scale (AIAS): A refined framework for educational assessment. Journal of University Teaching and Learning Practice, 22(7). https://doi.org/10.53761/rrm4y757

Raduescu, C. (2024). Perceived impact of generative AI on assessments: Comparing educator and student perspectives in Australia, Cyprus, and the United States. Computers and Education: Artificial Intelligence.

Sakirin, T., & Said, R. B. (2023). User preferences for ChatGPT-powered conversational interfaces versus traditional methods. Mesopotamian Journal of Computer Science, 22–28. https://doi.org/10.58496/MJCSC/2023/004

Stokkink, P. (2025). The Impact of AI on Educational Assessment: A Framework for Constructive Alignment. From https://doi.org/10.48550/arXiv.2506.23815

Styve, O., Virkki, T., & Naeem, U. (2024). Developing Critical Thinking Practices Interwoven with Generative AI Usage in an Introductory Programming Course. IEEE Global Engineering Education Conference (EDUCON), (pp. 1-8). Kos Island, Greece.

Verkuilena, A. (2025). Creating AI personas for educational settings: a PREPARED4ED approach. Proceedings of SEFI2025: 53rd annual Conference of the European Society for Engineering Education, “Engineering & Society” theme. Tampere, Finland.

Zhan, Y., Boud, D., Dawson, P., & Yan, Z. (2025). Generative artificial intelligence as an enabler of student feedback engagement: a framework. Higher Education Research & Development, 1289–1304.

BIOGRAPHICAL INFORMATION

Dr ***Jens Bennedsen*** is a Professor in engineering didactics. His research area includes educational methods, technology and curriculum development methodology. He has published more than 80 articles at leading education conferences and in journals. He is a member of the CDIO Council.

Reidar Lyng is Associate Professor at The Dept. of Physics at NTNU, and Expert in Educational Development with the Centre for Science and Engineering Education Development (SEED) at NTNU in Trondheim. He holds an MSc degree in Chemical Engineering and a PhD degree in Physical Chemistry. He has more than 30 years of education development from NTNU and several Swedish universities. His research and development interests are wide-ranging and include the systemic interplay between teachers, students, and learning spaces. Reidar is a CDIO regional co-leader for the EU.

Dr ***Siegfried Rouvrais-Delahaie*** is Associate Professor at IMT Atlantique, a ranked engineering technical university in France. His educational & research interests are in methods and tools for higher education transformations. He has published around 80 papers at leading higher education conferences and in journals. He was CDIO board member a decade ago.

Mika Suutari is the Head of Learning Environment Development at Turku UAS. He works at the Rector's Office and collaborates with the Vice Rector of Education, advancing the strategic development of education. In his role, he coordinates the utilisation of AI at Turku UAS and has been sharing best practices on both national and international stages. He holds a MSc degree in Applied Mathematics, has 26 years of experience at Turku UAS, and has previously served as a teacher, a supervisor in degree education, and a manager in the Learning Environment Services unit within corporate services.

Dr ***Joo Ghee Lim*** is Director of the School of Electrical & Electronics Engineering at Singapore Polytechnic, overseeing pre-employment training (PET) and continuing education and training (CET) programmes and leading curriculum innovation grounded in the CDIO framework. He holds a PhD in Computer Science & Engineering from the University of New South Wales (UNSW) and an MEng and BEng in Electrical Engineering from the National University of Singapore.

Mark Nivan Singh is Deputy Director of Educational Development at Singapore Polytechnic. Mark leads in the implementation of educational initiatives like CDIO, Flipped Learning, Self-Directed Learning and Pedagogy for the Professions. Part of his work also involves working closely with faculty to re-design their curriculum, innovate their teaching and learning approaches and conducting workshops and other professional development activities.

Corresponding author

Jens Bennedsen
Department of Electrical and Computer Engineering
Aarhus University
Finlandsgade 22,
DK-8200 Aarhus N
Denmark
jbb@ece.au.dk
+4541893090