



---

*Commentary*

## **Beyond assessment security: A critical policy analysis of four alternative strategies to uphold academic integrity and adopt the GenAI transformation of teaching and learning for an accredited engineering degree**

**Sasha Nikolic<sup>1,\*</sup>, Montserrat Ros<sup>1</sup>, Yasir M. Al-Abdeli<sup>2</sup> and Helen Fairweather<sup>3</sup>**

<sup>1</sup> Faculty of Engineering and Information Sciences, University of Wollongong, Wollongong, New South Wales, Australia

<sup>2</sup> School of Engineering, Edith Cowan University, Western Australia, Australia

<sup>3</sup> Head of Accreditation, Engineers Australia, Brisbane, Australia

\* **Correspondence:** Email: [sasha@uow.edu.au](mailto:sasha@uow.edu.au).

Academic Editor: William Guo

**Abstract:** Generative Artificial Intelligence (GenAI), exemplified by tools such as ChatGPT, has posed significant challenges and opportunities in the realm of academic integrity, particularly in engineering education. This commentary critically examines alternative strategies that go beyond traditional assessment security, aiming to uphold academic integrity while embracing the transformative potential of GenAI in teaching and learning. In the context of programs that rely on unit level outcomes to determine the overall student progression (not the programmatic approach to progression), this study identifies and explores four strategies that may offer potential improvements to assessment security: I - risk-level analysis, which aligns the mix of supervised and unsupervised assessments with their susceptibility to GenAI-enabled misconduct; II - adaptive grade scaling, which links marks for unsupervised work to performance in supervised tasks, thereby discouraging dishonest outsourcing; III - gatekeeper units, which embed high-stakes supervised checkpoints that verify the essential competencies before progression; and IV - maintaining the status quo, which exposes the limitations of solely relying on the existing security measures. The analysis highlights the potential benefits and weaknesses of each approach, supporting holistic decision-making on policies that improve the ability of students to meet the competency requirements (compared to

strategy IV) while fostering a culture of honesty and ethical behavior. By providing a thorough examination of these strategies, this study contributes valuable insights for educators, policymakers, and researchers, aiming to facilitate a balanced approach that aligns with the accreditation requirements and prepares students for the ethical use of GenAI in their professional careers. This commentary broadens and stimulates discussions on academic integrity in the GenAI era, thus providing practical guidance for educators and institutions.

**Keywords:** accreditation, assessment, engineering, Generative AI (GenAI), integrity, policy, STEM

## 1. Introduction

The rapid emergence of Generative Artificial Intelligence (GenAI) has disrupted foundational assumptions in higher education, particularly around authorship, assessment, and academic integrity. To frame the urgency and scope of this transformation, this introduction is structured around three core themes: the initial shock of GenAI's capabilities, the evolving role of accreditation, and the challenges it poses to traditional assessment models.

### 1.1. The Generative AI shockwave

In 2023, a wave of concern swept through universities worldwide with the revelation that ChatGPT-3.5, a cutting-edge Generative Artificial Intelligence technology, could pass exams. As a revolutionary advancement in AI, GenAI can autonomously create content such as text, images, and music, often mimicking human creativity and functioning as a personalized human assistant [1]. While the use of technology for cheating is not a new phenomenon [2,3], the advent of GenAI has significantly amplified the accessibility, detectability challenges, and risk associated with academic misconduct [4,5]. When using many traditional assessment approaches, this has made it difficult to determine exactly what contribution a student has made or even whether learning has taken place at all if GenAI was used.

While much of the concern has centered on text-based content, engineering education with a strong numerical focus is equally vulnerable, with research showing that ChatGPT-3.5 could facilitate students to pass many assessment types (e.g., quizzes, numerical problems, writing activities such as reflections), and its capability would only increase [6]. A repeat study in 2024 found that the capability with ChatGPT-4 did increase (being able to pass more assessment types), with a growing number of GenAI options available that can do more and be used more reliably and accurately to facilitate both cheating and learning [7]. This has resulted in many studies considering various assessment and academic integrity options [8–10]. Despite this, institutional guidance has largely focused on text-based tasks, with limited and often vague references to STEM-related activities such as coding and problem-solving [11].

The next generation of GenAI is increasingly becoming multimodal, allowing users to interact with files and generate audio, video, and images. As the diversity of the output widens, students have a greater capacity to engage with it to fulfil the needs for a greater variety of assessment tasks. This raises a critical question: Does the output generated by a computer truly represent a student's work? If not, submitting such content without an appropriate acknowledgement or referencing is generally

considered misconduct [12]. Yet the boundaries blur, especially when students craft the prompt, make evaluative judgments on the results, and iteratively refine the output. An analysis of top U.S. university policies revealed that concerns around GenAI are now central to academic integrity guidelines, particularly as institutions struggle to define authorship in a rapidly evolving technological landscape [13]. Recent university policies across Asia affirm that the undisclosed use of GenAI content constitutes academic dishonesty, even when the students are involved in prompt creation and refinement [14].

While critical analyses and systematic reviews continue to highlight GenAI's potential, particularly around the personalization of learning, these benefits are increasingly overshadowed by pressing concerns around the ethical, legal, assessment, and academic misconduct risks that it introduces into education [15–20]. These risks are reported as stemming from GenAI's human-like output that can evade plagiarism detection, its free and universal availability that lowers the barrier to large-scale unauthorized usage, its multimodal capacity to generate text, code, data, and images across almost every assessment format, and the equity gap it widens between students with differing levels of digital and prompt-engineering literacy or financial resources to use the best models. This has led to staff and student perceptions that are mixed, where they are appreciative of the potential benefits while concerned about the risks [21,22]. Addressing these perceived risks requires proactive policy action [20,21], including solutions that safeguard academic integrity. While the focus of this work is on academic integrity, it is important to acknowledge that GenAI integration is inevitable. As outlined in a recent Engineers Australia report, most engineers believe that GenAI use will become a critical skill [23]. As integration becomes embedded within the assessment design and learning environments, the competencies required of students will also evolve, and with them, the nature of academic integrity risks will shift.

## 1.2. Accreditation

Academic misconduct, unethical behavior, and falsely demonstrating competencies pose many risks, especially for institutions where quality assurance is vital to maintaining accreditation. This paper explores this topic through the lens of an accredited engineering degree. However, this analysis is highly relevant to any degree, accredited or not. The International Engineering Alliance (IEA) sets out the graduate attributes expected of graduate engineers under the Washington, Sydney, and Dublin Accords, to which Australia is a signatory. The graduate attributes are the basis for the entry-to-practice competencies developed by Engineers Australia, which also align with the Australian Quality Framework (AQF) levels: AQF 8, 7, and 6 for the Washington, Sydney, and Dublin Accords, respectively [24–26]. At present, there are sixteen (mandatory) elements of competency through which the three entry-to-practice professional competencies are demonstrated: Knowledge and Skill Base, Engineering and Application Ability, and Personal Attributes. At the degree level, or “programs” as they are termed in the US, the discipline is defined through the program's specification. This program specification is informed by the generic entry-to-practice competency elements, the associated attainment indicators, and industry and community expectations. The accreditation process assesses whether the program produces graduates who are fit for entry-to-practice in a clearly defined discipline. The program must be supported by a specification of the intended educational outcomes that incorporates the entry-to-practice standards. While this study focuses on engineering education, other accredited degrees share similarities (e.g., computer and

veterinary science, medicine and nursing).

Moreover, the education outcomes reflect the general dimensions for graduate expectations in the Australian Qualifications Framework (AQF), which was first introduced in 1995. The education outcomes need to be evidenced through an aggregate of assessment practices over a multitude of subjects or units (the term varies across institutions; the term units will be used for consistency).

### 1.3. The assessment challenge

While assessment selection is an important component of learning design, it is also considered by many as a line of defense to ensure academic integrity. The target of most focus relating to GenAI and evidencing competencies starts at unit-level assessment practices. A discussion of engineering assessment strategies can be found in [27]. Additionally, assessment security is an important part of the academic integrity process but should not override assessment validity [28]. Within engineering, a study by Nikolic, Sandison [7] outlined the short-term and long-term assessment security measures to help ensure academic integrity. For example, the study outlined how changes to the question design of online quizzes can impact the risk profile of use in relation to GenAI. A report undertaken for the Australian Government by Lodge, Howard [29] advised that assessment reform was needed, and that assessments should emphasize the following:

1. Appropriate and authentic engagement with AI;
2. A programmatic/systemic approach aligned with discipline and qualification values;
3. The process of learning;
4. Opportunities for students to work appropriately with each other and AI; and
5. Security at meaningful points across the program to inform decisions about progression and completion.

Even before GenAI, programmatic assessments were suggested as a pathway to optimize the assessment for learning and the assessment for decision-making [30]. The four strategies that will be discussed can integrate with such an approach, thereby prioritizing the learning process rather than the end product.

Key to this analysis is understanding that academic integrity is not a single-layer problem and cannot be secured by assessments alone. However, assessment has been the focus of discussion. An institutional-level framework presented by Ellis and Murdoch [3] highlighted the need for multiple strategies to work together. These multi-layered strategies, which were applied at different levels across an institution, correlate to the student's ability and attitude towards their learning and academic integrity. The course syllabi or outlines serve as one of these institutional levers, thereby directly communicating the expectations and practices to students. As highlighted by Ali, Collier [31], an analysis of 98 computing syllabi from U.S. R1 institutions (doctoral-granting universities with high research activity) found that most GenAI guidance was embedded within academic integrity policies, often restricting use due to concerns about accuracy, privacy, and its potential to hinder learning. However, the approaches varied substantially, with some instructors offering guidance on citations, naming specific tools, or framing GenAI as a collaborative assistant, thus illustrating how policy communication can reflect and shape institutional strategy.

A recent study by Corbin, Phillip [32] analyzed the various approaches that aim to communicate permissible AI use to students. This includes traffic light systems and declarative approaches, which

attempt to regulate behavior through structured rules and self-disclosures. However, these frameworks fundamentally misunderstand the challenge posed by GenAI. The distinction between structural and discursive changes highlights why such measures fail to ensure assessment validity. Relying on student compliance, they lack the means to monitor or enforce adherence, thus prohibiting what they cannot detect and directing behaviors they cannot verify. In AI-enabled education, these approaches offer the appearance of control without a real impact, underscoring the need for a more transformative assessment design.

To widen the lens on viable strategies and building on the need for multi-layered, institution-wide approaches, this commentary explores four specific options that engineering departments may be considering, each with potential benefits and drawbacks, as additional support mechanisms beyond traditional assessment security. The purpose of using engineering as a case study is to widen the debate and stimulate the academic community to discuss opportunities beyond assessments alone. Therefore, this study seeks to address the question: ‘What options are available to engineering departments to ensure the integrity of their degrees beyond assessment security?’ To address this question, this study attempts to identify and critically analyze the strengths and weaknesses of the various options implementable at the departmental level. This critical discussion is an important contribution to the field because exemplary solutions are unavailable, and most engineering departments are currently engaged in similar discussions without scholarly guidance. This work will help broaden the discussion and help with reflections.

## 2. Banning GenAI is not an option

Many initial reactions to the academic risks posed by GenAI were for educational institutions to ban it [33,34]. However, the consensus quickly changed and concluded that banning GenAI was not feasible [35]. Students would use it regardless, and how students used it quickly evolved [36]. This momentum has continued to gather steam as GenAI educational integration implementations grow and evolve [37]. There is growing recognition of the value it brings as a co-intelligence [1] or teammate [38]. Over time, it will continue to be embedded into our everyday devices and applications.

Beyond the many risks and drawbacks identified in the introduction, GenAI can positively benefit teaching and learning. Some examples include the following:

- Personalization of learning: providing personalized, enriching, and innovative ways to engage students and adapt content to best suit the learning needs of each student [39]
- Improved feedback: providing immediate feedback with a greater level of detail and personalization that enhances the learning support [40]
- Powerful information retrieval: the ability to provide information that rivals search engines and is capable of a fast and reliable summary of data tailored to the understanding of the individual [18]
- Support creativity and problem solving: providing students support to brainstorm ideas, probe into more complex solutions, and generate new multi-modal works [41]
- Streamlining of administrative tasks: providing support to academics to remove repetitive tasks, thus allowing them to provide a greater level of attention to teaching [42]

The only feasible option is learning to accept that GenAI will transform teaching and learning. To

do this, the academic community has been finding ways to reimagine assessments, thereby focusing on how to use GenAI tools for better assessments, developing and credentialing GenAI literacies, and focusing on human capabilities outside the realm of GenAI [43–45]. Additionally, the focus has been on finding ways to secure assessment practices [7,46] and communicate an acceptable use [32]. Fundamentally, the risk profile of an assessment is best associated with its delivery, which is conducted in either a supervised/secure or unsupervised/unsecured environment.

**Supervised/Secure Assessment (SA):** This is an assessment that provides some level of assurance that a student is undertaking the work and does not have access to unauthorized materials. This generally includes any face-to-face activities within some form of secure environment. This could include an exam in a hall, lecture theatre, tutorial class, laboratory, or secured computer lab. It is important to note that cheating is still possible via SA, a risk that remains underappreciated, largely due to a failure to anticipate the innovative strategies students might devise. Depending on the type, the amount of cheating should generally be limited and is dependent on many factors, including the risk/reward ratio [47].

Proctored online assessments, in which a student undertakes an online exam on their personal computer in an unsupervised environment but is supervised via software, is a grey area. While such software is recognized as rather accomplished and capable of detecting cheating, a growing area of research suggests that they can be compromised through technical backdoors or via alternate technologies such as deepfakes (manipulated videos or images) and contract cheating (outsourcing academic work); such software can be high risk [7,48–50].

**Unsupervised/Unsecured Assessment (UA):** This is an assessment that allows the student to undertake the work in their own time, where the teaching staff do not monitor who completes the work. A student can offload this work to a friend or contract cheater, engage in cheating collaboration websites, or use technologies such as paraphrasing tools or GenAI. Again, student use will generally be a function of the risk/reward ratio; however, it is unknown who undertook the work.

Just because an assessment is unsupervised does not necessarily mean cheating is easy. One scenario could be a take-home assessment followed by a viva to critically justify the solution methods and the outcomes. Another example could be context-driven, such as an assessment that specifically pertains to the discussion in a particular tutorial/laboratory class in which GenAI might not be trained. A second consideration is to explore assessment security options and what we can do to help prevent cheating from taking place. For example, if someone considered the risk of an online quiz being completed by GenAI, that online quiz can be seen in two ways. The first is to see it as high risk if all the questions are text-based or as low risk if they are based on figures and tables. This is a constantly moving target: it may be correct at the time of writing, but the technology is getting better at such a fast pace that trying to outrun it becomes a losing game. A detailed summary of long and short-term security options for engineering assessments is covered in Nikolic, Sandison [7].

### 3. Four potential strategies to improve assurance of competency

While assessment security is the first level of defense, institutions must consider multiple strategic layers and levers to ensure academic integrity by helping to support, advise, monitor, direct, and compel students [3]. To answer the research question, the team explored potential options that engineering departments could consider as a second level of defense, thus helping to ensure that students graduate with the competencies required for accreditation.



The options considered by the authors are summarized and presented in Table 1. The options were developed through the following:

1. **Brainstorming:** This was based on traditional techniques as well as applying co-intelligence by engaging in vigorous prompting sessions with ChatGPT-4o.

2. **Relevant expertise:** One of the authors is an engineering education expert focused on GenAI assessments and educational integration who completed extensive research on GenAI-based academic integrity. Two authors are Associate Deans of Education from two universities, which helped consider the faculty-wide perspectives. One author is an accreditation manager, who considers the impact of GenAI on accreditation processes and procedures.

The four strategies were developed through a collaborative process that combined structured brainstorming with targeted prompting using ChatGPT-4o to explore the possibilities, strengths, and weaknesses. Then, these ideas were critically shaped and refined through the authors' combined expertise in engineering education, academic integrity, faculty leadership, and accreditation. This integration of creative exploration and expert judgment helped to ensure that the strategies were grounded in practical, institution-wide realities.

3. **Community feedback:** Once developed, feedback was gained from the community via multiple presentations and webinars by the first author, including during the review process for continued refinement. It was first presented to a multi-institutional working party exploring GenAI impacts within engineering, organized by Engineers Australia, and then to multiple engineering schools across New South Wales. Each option was presented, and feedback via open discussion from the participants was used to refine the strengths and weaknesses of each of the options, and to consider if any option outside those presented was considered a possibility. These refinements were based on informal reflections by the authors on the themes and insights that emerged from these discussions.

The authors struggled to position the options within the literature, even with extensive searching and assistance from GenAI. The options were assessed based on their feasibility for implementation within engineering departments, their capacity to uphold academic integrity in the presence of GenAI, and their alignment with accreditation requirements, particularly those set by Engineers Australia. Each strategy was considered in terms of its pedagogical value, ethical implications, impact on student progression, and the degree to which it supports or complicates the program-level assurance of graduate competencies.

It is worth noting that the options in Table 1 assume a non-programmatic level approach to student progression. Whilst the implementation of programmatic approaches may very well be informed by some of these options, they can also introduce a different set of progression milestones and assessment strategies that are not possible with the traditional unit-based strategies for student progression.

**Table 1.** Summary of identified academic integrity options in the unit-based progression strategy.

No.	Option	Description	Advantages	Disadvantages
1	<b>Risk-level Analysis (low, mid, high)</b>	Categorize assessments by their cheating risk, tailoring supervision and controls accordingly.	Easy to analyze	Not all assessments fit the risk categories

1a	<b>Limit the impact of unsupervised assessments</b>	Shift weight to assessments that are secure and supervised, reducing GenAI-enabled cheating chances.	Easy to substitute assessment tasks from unsupervised to supervised	This may have workload implications and prevent the demonstration of ethical conduct when working unsupervised. Not all competencies may be easily validated using short-duration supervised assessments (e.g., effective group work).
1b	<b>Human Focus</b>	Emphasize human psychomotor and affective skills beyond AI capabilities	Project and laboratory work are already embedded.	Shifting learning outcomes focus would require a rethink of most current implementations. The capability of GenAI is a moving target
1c	<b>Embracing GenAI</b>	Assume and integrate AI use. Redesign learning activities and assessments.	Less risk management is required. The benchmark expectancy is raised.	Requires a redesign of learning and assessment so that the technology is factored into the student's contribution.
2	<b>Adaptive Grade Scaling</b>	Tie unsupervised assessment grades to supervised performance	No or minimal restriction on supervised weighting or types	Genuine students may be impacted if they struggle with supervised assessment components. It may increase pressure to perform on SA and may encourage cheating. Weighting can be subjective, and direct correlations (in learning outcomes) in both supervised and unsupervised may be challenging.
2a	<b>Discrepancy Scaling</b>	Reduce unsupervised marks when the gap with the supervised score exceeds a set range	Easy to implement.	There may not be a strong correlation between the competencies of SA and UA. Adjustments and strategy can be subjective.
2b	<b>Threshold Scaling</b>	Scale unsupervised marks only if the supervised score drops below the threshold.	Scaling occurs only above a threshold, fewer students may be impacted.	Same as 2a, and may be less secure. Adjustments and scaling can be subjective.
3	<b>Gatekeeper Units</b>	Engage with targeted high-stakes supervised checkpoint units that verify essential competencies before progression.	Little impact across most units.	Gatekeeper units place extra pressure on students, which may impact their ability to progress and complete them in a reasonable timeframe. Such units may also give students the wrong impression that it is okay to cheat in particular units.
4	<b>Leave it to Assessment Security</b>	Depend solely on enhanced assessment security and detection measures.	No changes to policy	Detecting cheating will become more difficult. There are significant risks that students may adjust their behavior to attain the highest marks, not necessarily the most actual learning. The unit coordinator remains responsible for identifying and implementing assessments with strong assessment security.



### 3.1. Option 1: A risk-level analysis (low, mid, high)

In Australia, most, if not all, universities consider a grade of 50 as the minimum threshold to suggest that the student has demonstrated the minimum required level of competency to pass a unit. Additionally, this may include other conditions (often referred to as hurdles) tied to mandatory assessment components (labs, end-of-semester exams). Internationally, passing thresholds above 50% are more common [31]. Some of the arguments for this threshold include the following:

- **Standardization:** A clear, uniform standard, such as a 50% passing grade, helps maintain consistency across different classes, schools, or educational systems. Additionally, this makes it easier to compare academic performance across different contexts.
- **Simplicity:** A 50% threshold is straightforward and easy to understand for students, parents, teachers, and administrators. It provides a clear goal for students who aim to pass a course.
- **Motivation:** Knowing that they need to achieve at least half of the available points can motivate students to at least reach this basic level of understanding of the course material.
- **Minimum Competency:** Setting the passing mark at 50% can be a compromise between being overly lenient and too stringent. It suggests a fair measure of a student's understanding of the material, which implies that they have a basic grasp of half the content or skills taught. Moreover, a 50% pass mark allows students who want to excel beyond the basic pass mark to be awarded levels of 'Credit', 'Distinction', or 'High Distinction' for additional or outstanding work.

Key arguments for shifting away from the status quo include:

- **Subjective Standard:** The 50% mark is somewhat arbitrary and may not accurately reflect a student's true understanding or competency. Different subjects and different types of assessments might require different thresholds to accurately measure competency.
- **Reduced Rigor:** A 50% minimum could be seen as setting the bar too low, potentially leading to a dilution of academic standards. It might encourage minimal effort, as students may do just enough work to meet the minimum requirement rather than striving for a deeper understanding.
- **Lack of Differentiation:** This threshold doesn't account for varying degrees of difficulty within the course content or between different courses. A pass of 50% in a highly challenging course might demonstrate more competency than the same grade in an easier course.
- **Impact on Learning Attitudes:** While a passing mark of 50% does not necessarily mean students only engage with half the content, it can sometimes contribute to a strategic or minimalist approach to learning. This may encourage a mindset focused more on meeting the minimum requirements, such as calculating the exact marks needed on a final exam to pass, rather than fostering a deeper engagement or intrinsic motivation to learn.

Considering that 50% is an acceptable standard, it is important to consider how this translates across a unit. Table 2 provides a simulated analysis by the authors (the process is provided in the header columns) of the risks through various ratios between supervised and unsupervised assessments. The table shows that if a student can use a resource such as GenAI to obtain 75% of the available unsupervised marks, then a student would only need to obtain 43.8% from the supervised component if the unsupervised components totaled 20%. If a student demonstrated only 43.8% of the

required competency, then is this acceptable? How much greater is the risk for a graduating student with 43.8%, 39.3%, or 33.3% competency as compared to a student with 50% competency? This analysis assumes there are no hurdle assessments or minimum thresholds applied to any assessment that could result in a Technical Failure. In reality, many units often impose a threshold hurdle that is mandatory to pass in one or more supervised assessments. If the pass threshold were raised from 50% to 60%, a unit with an unsupervised component (UC) of 30% would see the minimum required performance in the supervised component (AMP) increase from 39.29% to 53.57%. This shift enhances assessment integrity while allowing greater flexibility to include unsupervised assessments, and such an increase is worth debating.

A high-level risk assessment provides a structured opportunity for reflection, which allows institutions to identify potential vulnerabilities and take proactive steps to mitigate them. An example of such an institutional review can be found in [51], where a broad risk assessment was used to inform policy and practice at scale. At a more specific level, the School of Electrical, Computer, and Telecommunications Engineering at the University of Wollongong applied the risk analysis framework outlined in Table 2 to evaluate assessment-related risks across its units. This process was used to identify the priority units for review based on their exposure to academic integrity risks. Once high-risk units were identified, targeted mitigation strategies were considered to address the specific challenges of the units presented. Consideration included limiting the impact of UA, shifting assessment toward process over product, applying a greater human focus, or embracing GenAI.

**Table 2.** Assessment weight risk analysis with 50% pass threshold.

<b>Supervised Component (SC)</b>	<b>Unsupervised Component (UC)</b>	<b>Achieved Unsupervised Mark (AUM)</b> (assuming 75% of the mark is obtainable from GenAI or other means = $UC * 75\%$ )	<b>Mark Required (MR) in the supervised component to reach 50%</b> ( $50 - AUM$ )	<b>Assessment Mark required from supervised component to pass unit (AMP)</b> ( $MR / SC$ )
95.00%	5.00%	3.75%	46.25%	48.68%
90.00%	10.00%	7.50%	42.50%	47.22%
80.00%	20.00%	15.00%	35.00%	43.75%
70.00%	30.00%	22.50%	27.50%	39.29%
60.00%	40.00%	30.00%	20.00%	33.33%

### **3.1.1. Option 1a: Limit the impact of unsupervised assessment**

Based on the assumed 50% threshold to pass a unit and based on the analysis shown in Table 2, first thoughts may center on ensuring that all units adhere to a limit of a 10-20% unsupervised maximum. For many technical units, achieving a supervised assessment weighting of 80% may seem possible using supervised written or practical exams, class participation marks, demonstrations, presentations or interviews, and supervisor progress marks. One drawback of such an approach is that supervised assessments can increase the associated workload, and if unaccompanied by detailed marking rubrics, may also seem subjective, which may be difficult to implement if budgetary restrictions exist. Additionally, it may be possible to go too far and seek 100% supervised content;

however, this is unnecessary and unwanted because we need to put students in situations where they can demonstrate ethical conduct and work unsupervised.

Such a high threshold can be unrealistic for many research or project units with large amounts of unsupervised written work. Assessment security strategies can be employed to improve the risk. For example, greater weight can be given to a supervisor's evaluation based on students meeting weekly milestones and articulating their learning. However, the ratio may still not be suitable, so options are needed. Raising the pass threshold of a unit to 60%, as previously discussed, is another strategy that can support integrity while maintaining flexibility.

Limits can also be managed to some degree through a threshold hurdle. This implementation requires students to demonstrate a particular level of competency in a particular assessment to progress, regardless of the total number of marks achieved [52]. For example, a thesis project may require a student to obtain a threshold mark in the demonstration of their project. Its use can be somewhat aligned to cases where the competency is critical (e.g., assessment of communication competency); however, due to the assessment weighting compared to others, a student can pass the unit even if they fail that component, thus leaving the competency unverified [53]. Not passing a threshold hurdle could result in a technical fail condition or require the student to repeat attempts to pass the hurdle. It needs to be noted that threshold hurdles can add significant pressure on students, which can potentially lead to increased anxiety and stress, particularly if they are aware that failing to meet the threshold, which is arbitrarily set, could result in a fail or the need to retake assessments. Requiring students to retake assessments if they do not meet the threshold can be logistically challenging and resource-intensive for the students and faculty. It may also delay the students' progression in their studies.

### ***3.1.2. Option 1b: Human focus***

One option is to reimagine unsupervised assessments so that they focus on the human capabilities that extend beyond what can be done by GenAI or contract cheating [44]. This goes beyond education and into the workplace as humans navigate a potential skills relevancy shift in the decades to come. These human capabilities, such as psychomotor or affective skills, which extend beyond the competencies of GenAI and contract cheating, are essential to develop well-rounded professionals who can excel in the modern workforce. Engineering is well-positioned because project and laboratory learning are well established. A rebalancing is needed to complement the current focus on cognitive learning objectives with greater emphasis on the psychomotor and affective [54]. The work of Seery, Agustian [55] showcased how such competencies can be recorded, while the work of Dunne and Nikolic [56] outlined how the reimagination of assessment can bring out psychomotor objectives. A disadvantage of this approach is that engineering academics, by and large, have little expertise in confidently assessing these non-cognitive skills [57]. Additionally, there is the possibility of excluding certain students (e.g., those with physical disability for psychomotor skills).

### ***3.1.3. Option 1c: Embracing Generative AI***

The next option is to consider designing unsupervised activities so that they embed GenAI while acknowledging that the scope of AI collaboration can span several levels. This would require a substantial rethink of most current assessment designs; however, if GenAI is embedded, then the

assessment integrity risk is lowered. Having a 10-20% unsupervised limit is no longer needed. This removes the need for students to secretly use GenAI, thus opening the field and ensuring all students get a fair chance to use the same digital technologies towards their learning and assessment. One can move the needle of difficulty higher, which introduces more complex activities that require higher levels of evaluative judgement and cognitive thinking abilities [7,58]. Just as mathematics education removed many tedious activities with the integration of the calculator, moving students to more advanced concepts and topics required them to both use the calculator as a tool and know more than what the calculator could do. The advantage of GenAI comes from moving beyond enhanced learning experiences towards transformative ones [37].

Such policies are being implemented at some higher education providers using a two-lane approach [59]. The premise of such an approach centers on providing GenAI integration choices for supervised assessments and not prohibiting GenAI use for unsupervised assessments, that is, unsupervised assessments need to assume that GenAI may have been used in some form. The premise is that unsupervised assessments are used for learning, while supervised assessments are used to assess learning. Such a policy requires a rethink of many assessment implementations. This policy closely aligns with the long-term assessment security recommendations outlined in Nikolic, Sandison [7]. Additionally, the above could sit in the context of a shift from unit-based progression to that based on programmatic assessment regimes, in which a rethink of the course learning outcomes is needed. However, this is beyond the scope of this commentary.

### 3.2. Option 2: Adaptive grade scaling

A possible method to retain the use of higher levels of unsupervised assessments is through adaptive grade scaling. This allows for greater flexibility, provides students with opportunities to demonstrate the competencies of professional and ethical behavior for higher-stakes assessments, and demonstrates some level of trust between the institution and the student. The adaptive grading scale works by applying a predetermined system in which grades of unsupervised assessments may be adjusted based on the performance in supervised assessments to ensure that the grades reflect genuine understanding and learning. This would be used in cases where there should be a connection between the learning from the unsupervised assessments that are scaffolded and reassured in the supervised components.

Two design possibilities can arise in this case, both of which rely on subjective scaling decisions.

#### 3.2.1. Option 2a: Discrepancy scaling

This approach looks to identify differences between the unsupervised and supervised components. For example, if the grade difference between the UA and SA exceeds 15%, then the adaptive grading system activates. To do this, the percentage difference is calculated between the average UA and SA grades. Then, a sliding scale for grade adjustment is used based on the discrepancy percentage. The greater the discrepancy, the larger the potential adjustment and the bigger the impact of the subjective scaling factor. An example of a possible scale could be the following:

- **16-20% Discrepancy:** Reduce UA grade by 10%;
- **21-25% Discrepancy:** Reduce UA grade by 15%; and
- **>25% Discrepancy:** Reduce UA grade by up to 20% or trigger additional evaluation such as a viva voce.

### 3.2.2. Option 2b: Threshold scaling

This approach is similar to discrepancy scaling, except it is not triggered unless a student reaches a particular threshold. For example, if the threshold was 50% for SA, and the student scores 85% in UA and 50% in SA, then there is no change based on the 50% passing threshold. This recognizes that the student has met the minimum expected competency within the supervised assessment, thus expressing confidence that such competencies are reflected in the unsupervised assessment. However, if the SA mark is below the threshold, then scaling applies. For example, if a student scores 85% in UA and 45% in the SA, then the SA is below the threshold and scaling is applied. In this instance, the discrepancy is 40%. Once again, a possible scale could be the following:

- **16-20% Discrepancy:** Reduce UA grade by 10%;
- **21-25% Discrepancy:** Reduce UA grade by 15%; and
- **>25% Discrepancy:** Reduce UA grade by up to 20% or trigger additional evaluation such as a viva voce.

As with the other options, risks associated with such scaling apply. The first is recognizing that students may perform very differently across different assessment tasks. For example, students with some disabilities can be negatively impacted by various types of supervised assessments, which can be caused by increased anxiety and performance pressure. In such cases, mitigations such as reasonable accommodations, psycho-educational supports [60], or co-design [61] with vulnerable students can be considered.

This approach may also change the student's behavior, with students applying a greater focus on SA activity at the expense of UA activity and may adjust the risk/reward ratio to take a greater risk in cheating for SA. Additionally, there may be a problem with the feedback alignment. If grades are adjusted at the end of the teaching session, then the feedback on the coursework might not align with the final grade, thus confusing students about their actual performance and areas which need improvement. Finally, one of the biggest concerns of implementing such a system is the complexity it brings to grade calculation. This requires careful management and explanation. Misunderstandings or errors in the grade calculations could lead to grievances and appeals, thus increasing the administrative burden. Beyond this, a disadvantage to both these approaches (2a & 2b) is that students tend to be vehemently opposed to 'losing' marks that they believe they have honestly 'earned'. In our own experiences, we have seen students fight against late penalties despite them being clearly stated in unit outlines. This approach risks swamping the engineering department with appeals.

### 3.3. Option 3: Gatekeeper units

In engineering accreditation, a key outcome is that students can demonstrate attainment of all the competencies stipulated by the accreditation body. This is done at the course level and not at the individual level. It assumes that if a student can achieve at least 50% (or the required mark set by the institution/accrediting body), then a demonstration of attaining those competencies has occurred. However, as outlined in Option 1, the greater the percentage of UA at an assumed level of (unacknowledged GenAI use), the lower the probability that a student might not achieve all competencies. This is because the units contain a synthesis of objectives/competencies. Some of



those competencies might be demonstrated in the UA component and not the SA.

One way to remove the risk that every student doesn't obtain the required competencies is to implement gatekeeping units that focus on using SA. There are multiple methodologies to consider, each with its strengths and weaknesses. For example, in each academic term, a gatekeeper unit might be used to ensure that students have the minimum competencies required to scaffold to the next session. Alternatively, it could be used in one unit yearly, once every two years, or only in the final year. The advantage of such an approach is that it provides a greater confidence that the students can demonstrate a minimum standard of knowledge and skills in meeting all accreditation criteria across their degree. By requiring all students to pass these essential courses, educational institutions can maintain high standards across the program. These courses can motivate students to understand the unit matter better, knowing they must pass these key hurdles to advance in their studies. If feedback is applied correctly, then the gatekeeper subjects can help identify struggling students, which allows for timely interventions with additional support, tutoring, or advising.

However, as with the other scenarios, such an approach has weaknesses. In the absence of a shift from progression based on units, one possibility may be that gatekeeper units may give students the wrong impression that it's "ok to cheat in subject X but not subject Y". A key concern for students is an increased pressure and anxiety to pass such subjects, especially if it prevents them from advancing. Knowing that certain subjects are make-or-break can lead to negative health impacts and potentially encourage cheating or other forms of academic dishonesty due to the higher risk/reward ratio. This can have a substantial impact, such as the potential for high failure rates. Even more so, if a student has difficulty with specific learning or assessment methods, then it creates a barrier to progression. To eliminate such risks, it is imperative that a gatekeeper unit does not rely on just one assessment type but integrates a variety whilst upholding scrutiny to thoroughly ensure ethical assessment practices. Furthermore, this approach may make non-gatekeeper units appear less important in the students' minds, thus reducing their motivation. This is because assessments can play an important role in motivating students to learn [62].

### **3.4. Option 4: Leave it to assessment security**

Safeguarding academic integrity has always relied on a combination of systems and practices working together, and this must continue and improve going forward [3]. Until the recent uptake of GenAI tools, the default option for checking the authenticity of written assessments was to heavily rely on similarity reports generated using automated tools (noting that it did not prevent contract cheating risks). Whilst that approach may have served well in the era of non-GenAI to motivate students to act and learn with integrity, this approach no longer suffices, as there is growing evidence that doing nothing (else) will create risks to the profession and create confusion and distrust that will only escalate [4,6,17,63,64]. Many attempts have been made to communicate permissible AI use to students; however, these lack actual enforcement capabilities, thus creating an illusion of assessment security [32].

Additionally, there are significant risks to learning, as students end up adjusting their behavior to attain the highest marks, which does not necessarily result in any learning taking place. This is the greatest challenge that GenAI is forcing us to consider: how do we measure learning in the era of GenAI?

---

#### 4. Accreditation perspective

From an accreditation standpoint, particularly in the context of Engineers Australia, the assurance of graduate competency and ethical conduct is paramount in all units, whether gatekeeper or not. Accreditation bodies expect institutions to demonstrate, with evidence, that every graduate has met the required learning outcomes and possesses the professional and personal attributes expected of an engineering practitioner in one of the three occupational categories. This expectation has important implications for the strategies explored in this study.

A risk-level analysis strongly aligns with the accreditation requirements. It provides a defensible and auditable framework that links the assessment design to the potential for misconduct and enables engineering departments to demonstrate how risks are being systematically mitigated. By showing that supervised/secure and unsupervised/unsecured components are carefully calibrated based on their exposure to GenAI-related risks, this approach provides a level of transparency and rationale that provides accreditation panels with the confidence that graduate outcomes are attained by all students.

Moreover, adaptive grade scaling may be acceptable within an accreditation framework, provided it is implemented with care. Such a system would need to be underpinned by clear documentation, equity checks, and moderation processes that improve its objectivity. Accreditation bodies would expect that any scaling methodology be validated for fairness and reliability, alongside transparent communication to students about how their grades are derived. Without this level of assurance, there is a risk that the relationship between the assessment outcomes and the demonstrated competencies could be called into question, particularly if based on subjective criteria.

Gatekeeper units offer clear, high-stakes checkpoints where essential competencies can be verified under supervised conditions. This directly supports the accreditation requirement that programs ensure the development and verification of the full range of graduate attributes. However, the implementation of gatekeeper units must be carefully managed. Additionally, the occurrence of gatekeeper units should not negatively impact the need to demonstrate competency in other units and the expectation that ethical conduct is maintained throughout a program. Any competencies that are mapped to ‘non-gatekeeper’ units may be questioned by accreditation bodies or have the assurance of learning in those subjects challenged. Furthermore, high failure rates, student progression delays, or an unbalanced pressure could raise equity concerns, including the concentration of performance pressure in fewer, high-stakes assessments dispersed throughout the program. Institutions must be able to demonstrate that the appropriate support mechanisms are in place to ensure fair, reasonable, and robustly validated student progression.

In contrast, solely relying on traditional assessment security measures is unlikely to satisfy accreditation standards in the longer term. While necessary, assessment security alone does not provide the kind of program-level assurance that accrediting bodies seek, particularly when large portions of a program include unsupervised/unsecured assessments. Accreditation panels are increasingly expecting institutions to adopt a multi-layered approach that combines authentic learning and assessment design, institutional policies, and culture to uphold integrity. With this, there will also be a need for upcoming graduates to demonstrate competency in digital technologies that now underpin engineering and professional practice in the age of GenAI. This is an evolving landscape that is beyond the scope of this commentary.

From an assessment security perspective, several broader policy implications should also be



considered. Policies on GenAI use should be clearly published in student outlines and supported by robust moderation, appeal processes, and staff development initiatives. A recent review of such documentation highlighted that the current state is chaotic, is marked by significant inconsistencies and disparities, and underscores the urgent need for greater clarity and alignment with research-informed practices [31]. Additionally, institutions must demonstrate continuous improvement by gathering data on the effectiveness of their integrity strategies and using it to inform regular program reviews. Finally, accrediting bodies expect that major policy changes are made in consultation with key stakeholders, including students, staff, and industry representatives. In this regard, benchmarking plays a critical role in ensuring the whole sector improves, not only individual institutions.

Ultimately, accreditation requires more than compliance: it demands assurance. Institutions that adopt thoughtful, well-documented, and transparent academic integrity strategies will not only meet accreditation standards but will also strengthen the trust in the integrity of their graduates in an increasingly GenAI-mediated world.

## 5. Conclusions

In the context of Generative Artificial Intelligence, the exploration of alternative strategies to uphold academic integrity underscores the transformative impact of this evolving technology on teaching and learning. While the immediate focus in this commentary has been on engineering courses and the competency alignment with its accreditation requirements, the perspectives provided may still be considered valid in other disciplines. Integrating GenAI tools presents a multitude of challenges and opportunities in the educational landscape. These cannot all be considered in this paper. Additionally, the transition to progression based on program level achievement, rather than unit level achievement, has not been considered, as this affects the four options reviewed. The key findings and implications from this critical analysis into a set of plausible processes to improve the validity of assessment outcomes allows the academic community to reflect on the complexities of generating GenAI policies to help ensure integrity. The options canvassed in this work serve as a starting point for any review of assessment strategies aimed at improving the validity of the assessments. It is intended to initiate meaningful debate and stimulate the creation of innovative and holistic strategies.

Three key strategies were identified to possibly assist competency assurance: risk-level analysis, adaptive grade scaling, and gatekeeper units. Each strategy offers unique advantages and challenges, thus highlighting the need for a balanced approach that considers the specific context of engineering education. This allows the academic community to reflect on the available strategies, identify the strengths and weaknesses, and integrate these points into institutional policies that best reflect its values. The fourth strategy is now probably obsolete in any degree linked to domains of a professional practice and the robust demonstration of competency.

It is important to note that GenAI integration is only viable if the resource demands are acknowledged. Some strategies, such as limiting unsupervised assessments or introducing gatekeeper units, may increase staff workload through the need for more supervised assessments and increased invigilation. Moreover, shifting the focus from the final artefact of assessment to the learning process involved to arrive at it may significantly increase the staff workload; this is a challenge that could be mitigated through the strategic use of assessment strategies, technologies to support balance and

efficiency, and a possible shift to programmatic progression strategies that may embody a variety of assessment designs and learning activities. Whilst these latter issues have not been tackled in this commentary, for programs that continue to rely on progression based on unit-level outcomes, adaptive scaling systems may require technical development and staff training, especially if they are to be applied consistently and fairly. While some strategies, such as embracing GenAI in unsupervised tasks, can reduce the reliance on proctoring tools and align with a future-ready pedagogy, it may require significant investments in staff capability, infrastructure, and policy redesign. Scalability of each option will depend on the size of the cohort, workload capacity, institutional support, and the extent to which processes can be standardized across subjects or cohorts.

The potential educational impacts of these four strategies, and others, need to be considered and revisited over time as digital and AI technologies evolve. Approaches that promote authentic assessment designs and promote the uptake and ethical use of GenAI, such as integrating AI literacy into discipline contexts or those which refocus program learning outcomes to human capabilities, are more likely to motivate ethical assessment practices by students and positively engage them through the journey of deeper meaningful learning in the age of GenAI. Conversely, a heavy reliance on surveillance or punitive scaling could reduce trust, increase anxiety, or inadvertently shift the focus from learning to performance management. Therefore, a balanced implementation must consider not only the integrity outcomes but also student motivation, fairness, and the overall learning experience.

Institutions must implement multi-layered strategies that combine assessment security with more meaningful course outcomes that are aligned to employability as well as life-long learning. Regardless of the final mix of options selected, higher education institutions should foster a culture of excellence, competency building, equity, and ethical behavior in their programs. Students should see the value of demonstrating robust competency and relevant employability skills whilst they ethically practice and embody the social responsibilities expected in the contemporary workforce.

### **Author contributions**

Sasha Nikolic served as the project lead and principal author of this manuscript. Montserrat Ros, Yasir M. Al-Abdeli, and Helen Fairweather contributed their expertise through critical analysis of the concepts presented and by supporting the review process. They also played a key role in proofreading and ensuring clarity and consistency throughout the text. All authors reviewed the final manuscript thoroughly and contributed to its refinement.

### **Use of Generative-AI tools declaration**

As outlined in the methodology, ChatGPT-4o was used to support brainstorming and to explore the possibilities, strengths, and limitations of the ideas presented. Additionally, it assisted in reframing and refining the author's thoughts for clarity and conciseness, and in reviewing drafts to provide feedback that informed subsequent revisions. The SciSpace plugin was employed to help identify literature relevant to the four strategies discussed. All AI-generated content was critically reviewed and verified by the authors, who take full responsibility for the final manuscript. Additionally, Grammarly was used for proofreading to ensure accuracy and clarity. ChatGPT-o3 was trained in the STME referencing format and was used to convert the references created in EndNote, as no style template was available.

## Acknowledgments

We would like to thank the constructive feedback provided by the reviewers. This work is an initiative of the Australasian Artificial Intelligence in Engineering Education Centre (AAIEEC).

## Conflict of interest

The authors have no conflicts of interest in this paper.

Dr. Sasha Nikolic is an editorial member for STEM Education, and was not involved in the editorial review and the decision to publish this article.

## Ethics declaration

Human research ethics was not required for this work.

## References

1. Mollick, E., Co-intelligence: living and working with AI. 2024, London, UK: WH Allen.
2. Dawson, P., Five ways to hack and cheat with bring-your-own-device electronic examinations. *British Journal of Educational Technology*, 2016, 47(4): 592–600. <https://doi.org/10.1111/bjet.12246>
3. Ellis, C. and Murdoch, K., The educational integrity enforcement pyramid: a new framework for challenging and responding to student cheating. *Assessment & Evaluation in Higher Education*, 2024, 49(7): 924–934. <https://doi.org/10.1080/02602938.2024.2329167>
4. Cotton, D.R.E., Cotton, P.A. and Shipway, J.R., Chatting and cheating: ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 2023, 61(2): 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
5. Palmer, K., Sting operation fools a proctoring service—and results in blackmail attempt, in *Inside Higher Ed*. 2024. Available from: <https://www.insidehighered.com/news/students/academics/2024/03/28/sting-operation-fools-proctoring-service-blackmail-attempted>
6. Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G.M., Grundy, S., et al., ChatGPT versus engineering education assessment: a multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education*, 2023, 48(4): 559–614. <https://doi.org/10.1080/03043797.2023.2213169>
7. Nikolic, S., Sandison, C., Haque, R., Daniel, S., Grundy, S., Belkina, M., et al., ChatGPT, Copilot, Gemini, SciSpace and Wolfram versus higher education assessments: an updated multi-institutional study of the academic integrity impacts of generative artificial intelligence (GenAI) on assessment, teaching and learning in engineering. *Australasian Journal of Engineering Education*, 2024, 29(2): 126–153. <https://doi.org/10.1080/22054952.2024.2372154>
8. Lodge, J.M., de Barba, P. and Broadbent, J., Learning with generative artificial intelligence within a network of co-regulation. *Journal of University Teaching and Learning Practice*, 2023, 20(7): 1–10. <https://doi.org/10.53761/1.20.7.02>
9. Abeysekera, I., ChatGPT and academia on accounting assessments. *Journal of Open Innovation:*

- 
- Technology, Market, and Complexity*, 2024, 10(1): 100213. <https://www.sciencedirect.com/science/article/pii/S2199853124000076>.
10. Mustapha, K.B., Yap, E.H. and Abdalla Abakr, Y., Bard, ChatGPT and 3DGPT: a scientometric analysis of generative AI tools and assessment of implications for mechanical engineering education. *Interactive Technology and Smart Education*, 2024. <https://doi.org/10.1108/ITSE-10-2023-0198>
  11. McDonald, N., Johri, A., Ali, A. and Collier, A.H., Generative artificial intelligence in higher education: evidence from an analysis of institutional policies and guidelines. *Computers in Human Behavior: Artificial Humans*, 2025, 3: 100121. <https://doi.org/10.1016/j.chbah.2025.100121>
  12. Luo, J., A critical review of GenAI policies in higher education assessment: a call to reconsider the “originality” of students’ work. *Assessment & Evaluation in Higher Education*, 2024, 49(5): 651–664. <https://doi.org/10.1080/02602938.2024.2309963>
  13. An, Y., Yu, J.H. and James, S., Investigating the higher education institutions guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education*, 2025, 22(1). <https://doi.org/10.1186/s41239-025-00507-3>
  14. Dai, Y., Lai, S., Lim, C.P. and Liu, A., University policies on generative AI in Asia: promising practices, gaps, and future directions. *Journal of Asian Public Policy*, 2025: 1–22. <https://doi.org/10.1080/17516234.2024.2379070>
  15. Yu, H., Reflection on whether ChatGPT should be banned by academia from the perspective of education and teaching. *Frontiers in Psychology*, 2023, 14: 1181712. <https://doi.org/10.3389/fpsyg.2023.1181712>
  16. Jiang, Y., Xie, L., Lin, G. and Mo, F., Widen the debate: what is the academic community’s perception on ChatGPT? *Education and Information Technologies*, 2024, 29(1): 20181–20200. <https://doi.org/10.1007/s10639-024-12677-0>
  17. Ansari, A.N., Ahmad, S. and Bhutta, S.M., Mapping the global evidence around the use of ChatGPT in higher education: a systematic scoping review. *Education and Information Technologies*, 2024, 29(1): 11281–11321. <https://doi.org/10.1007/s10639-023-12223-4>
  18. Cong-Lem, N., Soyoo, A. and Tsering, D., A systematic review of the limitations and associated opportunities of ChatGPT. *International Journal of Human–Computer Interaction*, 2025, 41(7): 3851–3866. <https://doi.org/10.1080/10447318.2024.2344142>
  19. Quince, Z. and Nikolic, S., Student identification of the social, economic and environmental implications of using generative artificial intelligence (GenAI): identifying student ethical awareness of ChatGPT from a scaffolded multi-stage assessment. *European Journal of Engineering Education*, 2025, 1–20. <https://doi.org/10.1080/03043797.2025.2482830>
  20. Liu, Q., Hu, A., Gladman, T. and Gallagher, S., Eight months into reality: a scoping review of the application of ChatGPT in higher education teaching and learning. *Innovative Higher Education*, 2025, 1–24. <https://doi.org/10.1007/s10755-025-09790-4>
  21. Nikolic, S., Wentworth, I., Sheridan, L., Moss, S., Duursma, E., Jones, R.A., et al., A systematic literature review of attitudes, intentions and behaviours of teaching academics pertaining to AI and generative AI (GenAI) in higher education: an analysis of GenAI adoption using the UTAUT framework. *Australasian Journal of Educational Technology*, 2024, 40(6): 56–75.
-

- <https://doi.org/10.14742/ajet.9643>
22. Neupane, A., Shahi, T., Cowling, M. and Tanna, D., Threading the GenAI needle: unpacking the ups and downs of GenAI for higher education stakeholders. *Journal of Applied Learning and Teaching*, 2024, 7(2): 1–9. <https://doi.org/10.37074/jalt.2024.7.2>
  23. Bell, M., The impact of AI and generative technologies on the engineering profession. *Engineers Australia*, 2025. Available from: [https://www.engineersaustralia.org.au/sites/default/files/2025-01/impact-ai-generative-technologies-engineering-profession\\_0.pdf](https://www.engineersaustralia.org.au/sites/default/files/2025-01/impact-ai-generative-technologies-engineering-profession_0.pdf)
  24. Australian Qualifications Framework. 2016, Accessed 12 Feb 2016. <http://www.aqf.edu.au/>
  25. Engineers Australia, G02 Accreditation Criteria Guidelines. 2008, Accessed 25 July 2017. Weblink no longer active.
  26. Engineers Australia, Stage 1 competency standard for professional engineers. 2019, retrieved 07/02/25. <https://www.engineersaustralia.org.au/accreditation-management-system>
  27. Daniel, S., et al., Engineering assessment in the age of generative artificial intelligence: a critical analysis, in *2024 World Engineering Education Forum – Global Engineering Deans Council (WEEF-GEDC)*, Sydney, Australia, 2024.
  28. Dawson, P., Validity matters more than cheating. *Assessment & Evaluation in Higher Education*, 2024, 49(7): 1005–1016. <https://doi.org/10.1080/02602938.2024.2386662>
  29. Lodge, J., Howard, S. and Bearman, M., Assessment reform for the age of artificial intelligence. *Tertiary Education Quality and Standards Agency*, 2023. <https://www.teqsa.gov.au/guides-resources/resources/corporate-publications/assessment-reform-age-artificial-intelligence>
  30. van der Vleuten, C.P.M., Schuwirth, L.W.T., Driessen, E.W., Dijkstra, J., Tigelaar, D., Baartman, L.K.J. and Van Tartwijk, J., A model for programmatic assessment fit for purpose. *Medical Teacher*, 2012, 34(3): 205–214. <https://doi.org/10.3109/0142159X.2012.652239>
  31. Ali, A., Collier, A.H., Dewan, U., McDonald, N. and Johri, A., Analysis of generative AI policies in computing course syllabi, in *Proceedings of the 56th ACM Technical Symposium on Computer Science Education*. 2025.
  32. Corbin, T., Dawson, P. and Liu, D., Talk is cheap: why structural assessment changes are needed for a time of GenAI. *Assessment & Evaluation in Higher Education*, 2025, 1–11. <https://doi.org/10.1080/02602938.2025.2503964>
  33. Tlili, A., Shehata, B., Adarkwah, M.A., Bozkurt, A., Hickey, D.T., Huang, R. and Agyemang, B., What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 2023, 10(1): 15. <https://doi.org/10.1186/s40561-023-00237-x>
  34. Elkhodr, M., Gide, E., Wu, R. and Darwish, O., ICT students' perceptions towards ChatGPT: an experimental reflective lab analysis. *STEM Education*, 2023, 3(2): 70–88. <https://doi.org/10.3934/steme.2023006>
  35. van Dis, E.A.M., Bollen, J., Zuidema, W., van Rooij, R. and Bockting, C.L., ChatGPT: five priorities for research. *Nature*, 2023, 614(7947): 224–226.
  36. Purtill, J., ChatGPT was tipped to cause widespread cheating. Here's what students say happened, in *ABC Australia*. 2023. <https://www.abc.net.au/news/science/2023-11-22/how-high-school-students-used-chatgpt-2023-education-cheating/103108620>
  37. Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., et al., Implementing

- generative AI (GenAI) in higher education: a systematic review of case studies. *Computers and Education: Artificial Intelligence*, 2025, 100407. <https://doi.org/10.1016/j.caeai.2025.100407>
38. Fatahi, B., Nguyen, L.D., Khabbaz, H. and Hadgraft, R., Virtual teammates: transforming engineering learning through generative AI integration. In *35th Australasian Association of Engineering Education Conference*, Christchurch, New Zealand, 2024.
  39. Ruiz-Rojas, L.I., Acosta-Vargas, P., De-Moreta-Llovet, J. and Gonzalez-Rodriguez, M., Empowering education with generative artificial intelligence tools: approach with an instructional design matrix. *Sustainability*, 2023, 15(15): 11524. <https://doi.org/10.3390/su151511524>
  40. Barrett, A. and Pack, A., Not quite eye to AI: student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, 2023, 20(1): 1–24. <https://doi.org/10.1186/s41239-023-00427-0>
  41. Nikolic, S., Quince, Z., Lidfors Lindqvist, A., Neal, P., Grundy, S., Lim, M., et al., Project-work artificial intelligence integration framework (PAIIF): developing a CDIO-based framework for educational integration. *STEM Education*, 2025, 5(2): 310–332. <https://doi.org/10.3934/steme.2025016>
  42. Pinzolit, R., AI in academia: an overview of selected tools and their areas of application. *MAP Education and Humanities*, 2024, 4: 37–50. <https://doi.org/10.53880/2744-2373.2023.4.37>
  43. Bearman, M., Nieminen, J.H. and Ajjawi, R., Designing assessment in a digital world: an organising framework. *Assessment & Evaluation in Higher Education*, 2023, 48(3): 291–304. <https://doi.org/10.1080/02602938.2022.2069674>
  44. Bearman, M., Ajjawi, R., Boud, D., Tai, J. and Dawson, P., Cradle suggests... assessment and GenAI. *Centre for Research in Assessment and Digital Learning*, Deakin University, Melbourne, Australia, 2023. <https://doi.org/10.6084/m9.figshare.22494178>
  45. Belkina, M., Neal, P., Grundy, S., Hassan, G.M., Haque, R., Daniel, S., et al., Systematic literature review of GenAI integration in higher education and analysis of opportunities for engineering education. In *Proceedings of the 35th Annual Conference of the Australasian Association for Engineering Education (AAEE 2024)*, Christchurch, New Zealand. 2024, 1–10.
  46. Dawson, P., *Defending assessment security in a digital world: preventing e-cheating and supporting academic integrity in higher education*, 2020, New York, USA: Routledge.
  47. Noorbehbahani, F., Mohammadi, A. and Aminazadeh, M., A systematic review of research on cheating in online exams from 2010 to 2021. *Education and Information Technologies*, 2022, 27(6): 8413–8460. <https://doi.org/10.1007/s10639-022-10927-7>
  48. Burgess, B., Ginsberg, A., Felten, E.W. and Cohney, S., Watching the watchers: bias and vulnerability in remote proctoring software. In *31st USENIX Security Symposium (USENIX Security 22)*, Boston, MA, USA, 2022.
  49. Bergmans, L., Bouali, N., Luttikhuis, M. and Rensink, A., On the efficacy of online proctoring using Proctorio. In *13th International Conference on Computer Supported Education (CSEDU 2021)*, 2021.
  50. Dawson, P., Remote proctoring: understanding the debate, in *Second Handbook of Academic Integrity*, 2024, 1511–1526. Springer.
  51. University of Technology Sydney, Next steps for GenAI and assessment reform at UTS: a response to TEQSA. 2024.

- <https://educationexpress.uts.edu.au/blog/2024/09/02/next-steps-for-genai-and-assessment-reform-uts-response-teqsa/>
52. Hurdle requirements. 2025, accessed 19/05/2025. <https://www.adelaide.edu.au/learning/resources-for-educators/assessment/hurdle-requirements>
  53. GenAI and integrity: the arguments to reshape the thesis assessment structure. 2024, accessed 08/10/2025. <https://www.aaieec.org/post/genai-and-integrity-the-arguments-to-reshape-the-thesis-assessment-structure>.
  54. Nikolic, S., Suesse, T.F., Grundy, S., Haque, R., Lyden, S., Hassan, G.M., et al., Laboratory learning objectives: ranking objectives across the cognitive, psychomotor and affective domains within engineering. *European Journal of Engineering Education*, 2024, 49(4): 559–614. <https://doi.org/10.1080/03043797.2023.2248042>
  55. Seery, M.K., Agustian, H.Y., Doidge, E.D., Kucharski, M.M., O'Connor, H.M. and Price, A., Developing laboratory skills by incorporating peer-review and digital badges. *Chemistry Education Research and Practice*, 2017, 18(3): 403–419. <https://doi.org/10.1039/C7RP00003K>
  56. Dunne, I. and Nikolic, S., Autonomous assessment of a laboratory exam for the digital hardware curriculum. In *2021 IEEE International Conference on Engineering, Technology & Education (TALE)*, Wuhan, China, 2021.
  57. Nikolic, S., Suesse, T.F., Grundy, S., Haque, R., Lyden, S., Lal, S., et al., Assessment integrity and validity in the teaching laboratory: adapting to GenAI by developing an understanding of the verifiable learning objectives behind laboratory assessment selection. *European Journal of Engineering Education*, 2025, 1–28. <https://doi.org/10.1080/03043797.2025.2456944>
  58. Bearman, M., Tai, J., Dawson, P., Boud, D. and Ajjawi, R., Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 2024, 49(6): 893–905. <https://doi.org/10.1080/02602938.2024.2335321>
  59. Frequently asked questions about the two-lane approach to assessment in the age of AI. 2024, accessed 31/07/2024. <https://educational-innovation.sydney.edu.au/teaching@sydney/frequently-asked-questions-about-the-two-lane-approach-to-assessment-in-the-age-of-ai/>
  60. Yale-Soulière, G., Campeau, G., Turgeon, L. and Goulet, J., Evaluation of a brief intervention to reduce test anxiety in adolescents: a randomized control trial. *Current Psychology*, 2024, 43(39): 30760–30775. <https://doi.org/10.1007/s12144-024-06361-2>
  61. Smith, A., McConnell, L., Iyer, P., Allman-Farinelli, M. and Chen, J., Co-designing assessment tasks with students in tertiary education: a scoping review of the literature. *Assessment & Evaluation in Higher Education*, 2025, 50(2): 199–218. <https://doi.org/10.1080/02602938.2024.2376648>
  62. Leenknecht, M., Wijnia, L., Köhlen, M., Fryer, L., Rikers, R. and Loyens, S., Formative assessment as practice: the role of students' motivation. *Assessment & Evaluation in Higher Education*, 2021, 46(2): 236–255. <https://doi.org/10.1080/02602938.2020.1765228>
  63. Ferrara, E., GenAI against humanity: nefarious applications of generative artificial intelligence and large language models. *Journal of Computational Social Science*, 2024, 7: 549–569. <https://doi.org/10.1007/s42001-024-00250-1>
  64. Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z. and Trautsch, A., A large-scale comparison of human-written versus ChatGPT-generated essays. *Scientific Reports*, 2023, 13(1): 18617. <https://doi.org/10.1038/s41598-023-45644-9>



## Author's biography

**Dr Sasha Nikolic** is a Senior Lecturer at the University of Wollongong with a PhD in Engineering Education. He has worked in academia and industry. Dr. Nikolic has received multiple teaching and learning awards, including Australian Awards for University Teaching in 2012 and 2019, and AAEE awards in 2019 and 2023. He served in multiple governance roles with IEEE, and is an Associate Editor of the European Journal of Engineering Education. He currently serves as President for both the Australasian Association of Engineering Education and the Australasian Artificial Intelligence in Engineering Education Centre, where he leads multi-institutional initiatives on Generative AI.

**Dr Montserrat Ros** is an Associate Professor in the School of Electrical, Computer & Telecommunications Engineering and Associate Dean Education in the Faculty of Engineering and Information Sciences at the University of Wollongong, Australia. Montse graduated from a BE(Comp. Sys)/BSc(Math) with First Class Honours in 2000 and obtained her Ph.D. in Computer Engineering in 2007; both from the University of Queensland, Australia. Her research interests include Embedded Systems, Internet of Things, Sensor Networks Data Fusion, Code Compression and Engineering Education and has published over 100 peer-reviewed papers. Montse was awarded a 2019 Citation for Outstanding Contributions to Student Learning at the Australian Awards for University Teaching; and was named a UOW 2016 Woman of Impact for inspiring the young STEM generation.

**Dr Yasir M. Al-Abdeli** (Ph.D. USyd 2004) is an Associate Dean Teaching and Learning in a School of Engineering. Whilst his main areas of applied research are Thermofluids and energy, he has published on a number of topics within engineering education, including flipped classes and problem-based learning. He has a passion for continuous change in higher education enabled by technological change coupled, sound curriculum design, and good pedagogy so as to support learners in their lifelong learning. He is the co-founder of a community of practice on blended learning in Perth, Western Australia.

**Dr Helen Fairweather** is Head of Accreditation at Engineers Australia. She has over a decade of experience as an environmental engineering academic at the University of the Sunshine Coast, where she integrated professional and academic experience into the engineering curriculum during a period of significant societal change. Dr Fairweather now applies this expertise to influence the engineering profession at a national level. Her recent work has focused on promoting diversity in engineering education and highlighting the critical role of engineers in advancing the United Nations Sustainable Development Goals.



AIMS Press

©2025 the Author(s), licensee by AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>).