ADAPTING ASSESSMENT IN THE GENAI ERA: AN UNDERGRADUATE PHYSICS CASE STUDY

Quince, Z.[a,b], Faulconer, E.[c]

Presenting Author: Emily Faulconer (emily.faulconer@monash.edu)
[a]Faculty of Science and Engineering, Southern Cross University, Coolangatta, Queensland, 4225, Australia
[b]Centre for Teaching and Learning, Southern Cross University, Coolangatta, Queensland, 4225, Australia
[c]School of Biological Sciences, Monash University, Clayton, Victoria, 3800, Australia

KEYWORDS: Generative AI (GenAI), Higher Education, Academic Integrity

PROBLEM
The widespread adoption of Generative AI (GenAI) tools in higher education, while offering learning opportunities, can also compromise academic integrity and assessment authenticity. This necessitates a proactive approach with clear institutional policies to guide appropriate use of GenAI.

PLAN
Assessments for a 2nd year physics unit were evaluated for vulnerability to GenAI misuse to determine the appropriate level of use according to institutional policies, as well as evaluating the need for redesign to ensure authenticity and resistance to GenAI misuse.

ACTION
A risk matrix was applied to assessment tasks to examine the likelihood of misuse and potential impact on learning, informing GenAI use category designation and flagging highly vulnerable tasks for redesign. Adopting GenAI use categories maintained constructive alignment and reduced opportunities for inappropriate GenAI use. Clear guidance on acceptable use was embedded in briefs and rubrics, supported with workshops and educational tools like infographics.

REFLECTION
The structured assessment evaluation using the AAM-GenAI led to clear use designation and informed targeted redesign, where constructive alignment and a risk audit were undertaken. A key finding is that adoption of new or adjusted institutional guidelines for GenAI use has a non-trivial workload and may trigger larger collaborative projects to ensure GenAI resistant assessments, ensuring students achieve the learning outcomes. Future work will explore the impact of varying GenAI use on assessment tasks on their performance, including quantifying unethical use.

INTRODUCTION
Generative artificial intelligence (GenAI) tools such as ChatGPT have transformed how higher education STEM students engage with activities, content, and assessments (e.g., Wang et al., 2024). While these technologies offer powerful opportunities for support and creativity, they also pose new challenges to academic integrity (e.g., Quince and Nikolic, 2025), assessment authenticity (Powell & Forsyth, 2024), and curriculum design (Dickey & Bejarano, 2024).

Globally, institutions have responded to this emerging technology by developing policies for its use (Wang et al., 2024). Southern Cross University (SCU) implemented GenAI Tool Use Descriptors, outlining permissible levels of AI tool use in assessments across three categories: Full Use, Purpose-Specific Use, and No Use (Quince and Greenaway, 2025), mandating assessment evaluation and alignment with these descriptors to ensure clarity, fairness, and integrity in evaluation.

This paper presents a case study of Physics II: Electricity and Circuits (PHYS2001), a 2nd year undergraduate physics unit at SCU. We detail the process of reviewing assessment tasks in response to the GenAI guidelines and ensuring the assessment were GenAI resistive. Our goal was to align each task with the appropriate GenAI use category while ensuring continued alignment with the unit's four Unit Learning Outcomes (ULOs). This work contributes to the broader conversation around assessment redesign in STEM disciplines in the era of GenAI enabled learning.

BACKGROUND AND LITERATURE REVIEW

Constructive Alignment in Assessment

The constructive alignment (Biggs,1996) *framework for curriculum design aligns learning outcomes, teaching methods, and assessment, ensuring that students are engaging deeply in order to attain the intended learning outcomes*. This framework has been applied in STEM disciplines (e.g., Cole et al., 2019), where performance-based and applied knowledge is central (Reeves, 2006). Assessment redesign must preserve the link between what is taught and how students demonstrate mastery. When assessments drift from intended learning outcomes, particularly with new technologies in play, the risk increases that students may succeed in the task but fail to develop and display the intended conceptual understanding or practical skills.

GenAI and Assessment Integrity

GenAI tools are increasingly used by students, at times without awareness of academic integrity boundaries and writing ethics (e.g., Oh et al, 2023; Rasul et al., 2024). Institutions are grappling with how to embrace these tools without compromising learning. Without clear expectations, students may use GenAI in ways that shortcut the learning process by generating output disengaged from the student’s own reasoning or knowledge construction. SCU’s policy initiative aims to bridge this gap. GenAI Tool Use Descriptors clarify expectations to make acceptable GenAI use explicit, safeguarding against unintentional misconduct while allowing appropriate innovation and development of digital literacy skills (Quince and Greenaway, 2025).

SCU’s Institutional Approach

SCU’s Centre for Teaching & Learning has published a GenAI strategy, including the Assessment Adaptation Model (AAM-GenAI) which guides academics through evaluating GenAI vulnerability of tasks, determining the desired level of tool use, and modifying assessments accordingly (Quince et al., 2025). It encourages educators to revisit the question: does this assessment still demonstrate the intended learning outcome if completed with the assistance of GenAI? If not, how should it be redesigned, or is the assessment no longer fit for purpose?

CURRICULUM REVIEW PROCESS

Unit Context

PHYS2001 supports the Bachelor of Engineering, Bachelor of Education, and related programs as a core unit. It introduces electromagnetism, circuit theory, and experimental physics while developing student capabilities in analysis, experimental design, and scientific communication. Enrollment ranges between 25 to 40 students per offering, drawing primarily from two distinct cohorts: pre-service science teachers and engineering undergraduates. This interdisciplinary cohort influences content delivery and pedagogical design, with an emphasis on practical application and conceptual clarity suited to future educators and technical professionals alike.

The assessment structure reflects the applied nature of physics, progressively building skills. Initial review revealed that the assessments were constructively aligned with the unit’s four learning outcomes (Table 1). Assessment briefs did not contain any references to GenAI use or disclosure of use nor did existing rubrics account for or assess responsible technology use.

Table 1: Constructive Alignment Assessment of Learning in PHYS2001

ULO	Assessment	Module	Teaching & Learning Activities
1. Apply foundational concepts in electricity and circuits, including the application of Coulomb's law, Ohm's law, Kirchhoff's voltage law, and Kirchhoff's current law, to analyse and solve electrical circuit problems effectively	Laboratory Workbook (A2) Quiz (A3)	1: Electrostatics 2: Electric circuits	•Lectures •Tutorials •Practice quizzes •Worked example videos
2. Effectively apply knowledge of electricity and circuits to solve simple industry-related problems	Oral Presentation (A1)	2: Electric circuits 3: Electronics	•Case-based learning •Guest content developed by industry •Workshops on applications •Discussion forums

3. Examine how circuit properties influence trends in semiconductor tech and applications	Oral Presentation (A1)	2: Electric circuits 3: Electronics	•Lectures •Readings on current technologies •Discussion activities •Research-based tasks
4. Perform, analyse, and report on experiments involving circuits, semiconductors, and transistors	Laboratory Workbook (A2)	1: Electrostatics 2: Electric circuits 3: Electronics 4: Electronic systems 5: Magnetism 6: Electromagnetism	• Simulated Laboratory sessions • Demonstrations • Lab preparation guides • Post-lab debriefs

Design Plan

To implement the new institutional policy, the unit assessor (UA) initiated a review of assessment briefs, rubrics, and student-facing guidance to ensure transparency, support ethical decision-making, and maintain alignment with learning outcomes. GenAI Tool Use Descriptors were assessed by the UA for alignment with the intended learning, and assessment briefs were updated to explicitly state the category (1 - Full Use, 2 - Purpose-Specific Use, and 3 - No Use) in order to provide clear guidance on permitted and prohibited uses of GenAI tools for each assessment type (Table 2). Time and workload capacity required a top-down approach where the ULOs were unable to be changed.

Table 2: Assessment Specifications, Including Evolution of GenAI Policy

Assessment Type	Weight (%)	Purpose	GenAI Use Policy (Before)	GenAI Use Policy (After)
A1 – Oral Presentation	20	Synthesis and communication of scientific content	No Use	Purpose-Specific
A2 – Laboratory Workbook	50	Analysis and interpretation of experimental data and written communication of findings	No Use	Full use
A3 – Quiz	30	Reinforcement of key theoretical concepts and problem-solving techniques	No Use	Full use

While applying Tool Use Descriptors may appear to be a minor task, their integration required more than insurance of learners' compliance; it necessitated a time-intensive, collaborative, and pedagogically-informed design process that preserved the integrity, fairness, and learning value of existing tasks. The goal was to make assessment expectations explicit, supporting ethical engagement with emerging technologies, and ensuring students were still authentically assessed against the unit's learning outcomes, without engaging in a large-scale redesign. To achieve this, the UA adopted the Assessment Adaptation Model for Generative AI (AAM-GenAI) as a guiding framework (Figure 1). Due to time constraints the UA commenced with Stage 2, Analyse, using existing assessments as the starting point.

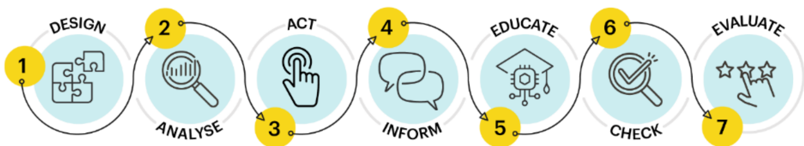

Figure 1: Assessment Adaptation Model for GenAI (Quince et al., 2025)

AAM-GenAI Step 2: Analyse

A systematic audit of each assessment task was performed using a risk matrix to identify the likelihood of GenAI misuse and potential impact of inappropriate use on student learning (Table 3). Likelihood of use was considered ranging from rare (almost no realistic misuse) to possible (some exploitation with varying skill) to almost certain (easy exploitation of GenAI). Impact on learning was considered ranging from negligible (no discernible impact with students still developing core skills and knowledge) to moderate (significant learning is undermined) to critical (the entire purpose of the assessment is nullified).

Table 3: Assessment Task Risk Matrix

		Impact on learning				
Likelihood of		*Negligible*	*Minor*	*Moderate*	*Major*	*Critical*
	Almost Certain	Low-Medium	Medium	High	High	Critical
	Likely	Low	Low-Medium	Medium	High	High
	Possible	Low	Low-Medium	Medium	Medium-High	High
	Unlikely	Low	Low	Low-Medium	Medium	Medium-High
	Rare	Low	Low	Low-Medium	Medium	Medium

A1 – Oral Presentation

The slide deck for the oral presentation could be quickly prepared by GenAI. If a student is generating narrative content using GenAI without verifying accuracy and critically analysing the output, it would be possible to create a final product without demonstrating an understanding of the content at the level of mastery expected based on the aligned learning objectives. Though less likely, students with strong GenAI skills may even be able to generate a deepfake, a digitally created video representation of themselves. Because even the most basic use of GenAI for this task could shortcut demonstration of learning, the researchers determined that the likelihood of use is “likely” (4) and the impact to learning is “major” (4). This sets a risk rating of high, warranting an intervention to evaluate the ideal GenAI use category and update instructions and rubric accordingly.

A2 – Laboratory Workbook

The laboratory workbook could be prepared by GenAI, though a large time investment would be needed to match the specific parameters. This would likely take more time than to undertake the experiments and complete the workbook. However, some workbook questions may be vulnerable to generation of content by GenAI. The risk of easily retrievable data to feed into a GenAI prompt may be lower than publicly available activities because these were designed by the unit teaching team. Because of the ability to use GenAI for multiple sections and because of the generally compromised nature of data sets for introductory physics experiments through websites like CourseHero and programs like Discord, the researchers determined that the likelihood of use is “likely (4) and the impact to learning could vary between “moderate” (3) to “major”, depending on whether the student used their generated data or whether they scraped data from an online source.

A3 –Quiz

We found online quiz questions were compromised by GenAI with 100% accuracy with a short time investment. As the questions were multiple choice, it increases the GenAI tool’s ability to get the correct answers (Nikolic et al., 2023). Further, there were no variations in the questions between students (question pools) and no requirement to upload solution calculations. Given the assessment format and the vulnerability to GenAI use to obtain accurate responses, the likelihood was deemed to be “almost certain” (5), with “critical” (5) impact as students using GenAI could earn high marks without demonstrating content knowledge or engaging in higher order thinking. Combined, this assessment places this assessment task in the critical-risk category, warranting intervention through assessment redesign that goes *beyond* an update to instructions and rubric.

AAM-GenAI Step 3: Act

Based on the analysis, assessment tasks were designated a GenAI use categories (Table 1) and subsequent steps to update the assessment briefs and rubrics followed. The goal was not to create a “GenAI proof” assessment but to realign with authentic, narrative driven assessment that pushes students to critically and creatively engage. Edits emphasized ethical use of the technology and all task instructions were updated to require an acknowledgement of GenAI use, as follows:

> “I acknowledge that I [**did / did not**] use Generative AI tools. in the completion of this assignment. If used, the tool(s) I used were [X]. I used them for the purpose(s) of [X] I confirm that any use of AI was, ethical, and in support of my own learning, and that the final submission reflects my original understanding, effort and is demonstrably my own work.”

The vulnerability of each assessment was re-evalauted as a result of these edits (Table 4).

Table 4: Assessment Task GenAI Vulnerability

Assessment	Risk Category (Before)	Risk Category (After)
A1	High	Low-Medium (Possible + Minor)
A2	Medium to High	Medium (Possible + Moderate)
A3	Critical	*Medium* (Likely + Moderate)

A1 – Oral Presentation
The oral presentation was designated Purpose-Specific Use, allowing for background research, idea generation, or drafting, but not for verbatim narrative content or to mimic speech. This aligns with engineering professional practice. Misuse can be associated by formatting (e.g., American spelling and bold key words) and Turnitin's AI rating. Whilst these may provide indicators, they should not be used in isolation. The key to an engaging presentation is good use of imagery to support the verbal communication. Even good GenAI presentation platforms (e.g.,Canva) require input to create acceptable flow and images. Students were actively encouraged to use GenAI to help draft their speeches and slide design and guided to what inappropriate use was i.e. generating a speech and slide deck with no modifications or editing the flow and content.

Permissible use of GenAI can enhance learning, especially when students are guided in using it to support learning outcomes and its use is communicated as ethical and aligned with social norms (Hamerman, 2024). The emphasis remains on students' personal understanding and ability to communicate physics concepts clearly and authentically. By embedding GenAI into the task, the likelihood of misuse may have been lowered, though educators and students have slightly misaligned perspectives on acceptable GenAI use in the writing process, which warrants explicit guidelines in instructions (Barret & Pack, 2023). Given the adjustment to support certain use of GenAI on the assessment task, the likelihood of inappropriate use was hypothesized by the researchers to drop from "likely" to "possible" and the impact on learning dropped to minor.

A2 – Laboratory Workbook
The Laboratory Workbook was designated Open Use, which permits use of GenAI tools for all tasks. This again mimics engineering professional practice where professionals have access to tools, best practice examples and colleagues to help solve the problems. While this alignment with professional practice is pedagogically valuable, we acknowledge that students may require more structured support to use such tools effectively. The instructions and rubric for the laboratory reports were updated to require documented evidence of completion of the online simulations (screenshots and show-work calculations) and a short weekly reflection. It has been demonstrated that a low effort assessment task carries less weight with students misusing GenAI tools (Toker and Akgun, 2024). This implementation of guided use, and enhanced critical reflection aligned with content and examples discussed in the units content. This is not a GenAI proof assessment, but it becomes more resistant with the appropriate guidance and scaffolding. With the adjustment to the use of GenAI on the assessment task, the likelihood of inappropriate use was hypothesized by the researchers to drop from "likely" to "possible" and the impact on learning dropped to moderate.

A3 –Quiz
The Quiz was designated Open Use, mimicking the ability in the workplace to readily turn to tech tools to fill in your knowledge gaps. Allowable use of GenAI on summative assessments is an emerging area of research with unclear findings at this time. As this assessment was critically vulnerable to GenAI misuse, it warranted a redesign beyond the minor edits to instructions and rubrics for the previous assessment tasks. The nature of the redesign is beyond the scope of this work. The new risk category for this assessment was initially pending determination as it underwent collaborative evaluation by students. Since the full implementation, this risk was updated to a likely likelihood of misuse and a moderate impact on learning..

AAM-GenAI – Step 4: Inform
Once the GenAI use levels were determined (Step 2), assessment briefs and rubrics were revised (Step 3). To inform students of these changes, the UA provided examples of acceptable and unacceptable applications of GenAI within that task, provided regular verbal and written reminders of students' responsibility to acknowledge any use of such tools, and embedded referencing to SCU's broader policy on GenAI and academic integrity in multiple ways (learning management system announcements, etc.). A "GenAI in PHYS2001" infographic was also uploaded to the learning site summarising the allowed usage for each task, and students were encouraged to seek clarification prior to submission if unsure.

AAM-GenAI – Step 5: Educate
To support student uptake and ensure clarity, GenAI expectations were introduced early in the semester. The first workshop included an overview of SCU's GenAI policy and how it applied specifically to PHYS2001. Example scenarios were discussed to help students distinguish between supportive and inappropriate uses of AI tools.

AAM-GenAI – Steps 6 (Check) and 7 (Evaluate)
While formal processes for Stages 6 and 7 have not yet been fully implemented, the UA continues to monitor for potential GenAI misuse. Informal student and tutor feedback is being gathered throughout the semester to evaluate the clarity, fairness, and impact of the Gen-AI use in the assessment tasks. Insights from this implementation will inform future evaluation cycles to strengthen both compliance and constructive student learning experiences. Planned activities include an anonymous survey following the oral presentations, reflective commentary prompts built into selected submissions, and informal tutor debriefs to capture insights into how students are engaging with the revised assessment expectations.

CONCLUSION

After confirming constructive alignment, the vulnerability of assessments were evaluated using the AAM-GenAI model, applying a risk matrix approach. This supported identification of Tool Use labels aligned with institutional policy and informed targeted edits to instructions and rubrics. Furthermore, one particularly vulnerable assessment was flagged for more intensive redesign. This implementation is not a one-time policy response, but the first stage of an iterative design cycle. By foregrounding student engagement, ethical reasoning, and authentic demonstration of learning, the process reflects a commitment to educational integrity that evolves with technological change. This case study affirms the value of approaching curriculum innovation not as a compliance exercise, but as a pedagogically grounded opportunity to clarify values, build student trust, and prepare learners for a digitally mediated academic and professional future. As new insights emerge, they will contribute to the continuing refinement of PHYS2001, and potentially, to wider conversations about how STEM educators can respond to AI with clarity, flexibility, and purpose.

REFERENCES

Almusharraf, N., & Bailey, D. (2023). Exploring students' perceptions of ChatGPT in higher education: Experiences, benefits, and concerns. *International Journal of Educational Technology in Higher Education, 20*(1), Article 52. https://doi.org/10.1186/s41239-023-00427-0

Barrett, A., & Pack, A. (2023). Not quite eye to A.I.: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education, 20*(59). https://doi.org/10.1186/s41239-023-00427-0

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*(3), 347–364. https://doi.org/10.1007/BF00138871

Cole, R., Reynders, G., Ruder, S., Stanford, C., & Lantz, J. (2019). Constructive alignment beyond content: Assessing professional skills in student group interactions and written work. In M. Schultz, S. Schmid, & G. Lawrie (Eds.), *Research and practice in chemistry education* (pp. 173–188). Springer. https://doi.org/10.1007/978-981-13-6998-8_13

Dickey, E., & Bejarano, A. (2024). GAIDE: A framework for using generative AI to assist in course content development. *2024 IEEE Frontiers in Education Conference (FIE)*, 1–9. https://doi.org/10.1109/FIE61694.2024.10893132

Hamerman, E. J., Aggarwal, A., & Martins, C. (2025). An investigation of generative AI in the classroom and its implications for university policy. *Quality Assurance in Education, 33*(2), 253–266. https://doi.org/10.1108/QAE-08-2024-0149

Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., Lyden, S., Neal, P., & Sandison, C. (2023). *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, *48*(4), 559–614. https://doi.org/10.1080/03043797.2023.2197388

Oh, N., Choi, G. S., & Lee, W. Y. (2023). ChatGPT goes to the operating room: evaluating GPT-4 performance and its potential in surgical education and training in the era of large language models. *Annals of Surgical Treatment and Research*, *104*(5), 269-273. 10.4174/astr.2023.104.5.269

Powell, S., & Forsyth, R. (2024). Generative AI and the implications for authentic assessment. In *Using generative AI effectively in higher education* (1st ed). Routledge. https://doi.org/10.4324/9781003482918

Quince, Z., & Greenaway, R. (2025). *Practice report: GenAI policy evolution at Southern Cross University* (Southern Cross University Scholarship of Learning and Teaching Paper No. 26). SSRN. https://doi.org/10.2139/ssrn.5330009

Quince, Z., Munn, J., & Greenaway, R. (in press). *Practice report: Adapting assessment in the age of generative AI: The AAM-GenAI framework*. Manuscript accepted for publication, Southern Cross University Scholarship of Learning and Teaching Papers.

Quince, Z., & Nikolic, S. (2025). 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*, 1-20.https://doi.org/10.1080/03043797.2025.2482830

Rasul, T., Nair, S., Kalendra, D., Balaji, M. S., de Oliveira Santini, F., Ladeira, W. J., Rather, R. A., Yasin, N., Rodriguez, R. V., Kokkalis, P., Murad, M. W., & Hossain, M. U. (2024). Enhancing academic integrity among students in GenAI era: A holistic framework. *The International Journal of Management Education, 22*(3), 101041. https://doi.org/10.1016/j.ijme.2024.101041

Reeves, D. B. (2006). *The learning leader: How to focus school improvement for better results.* ASCD.

Toker, S., & Akgun, M. (2024). *The role of task complexity in reducing AI plagiarism: A study of generative AI tools* [Preprint]. arXiv. https://arxiv.org/abs/2412.13412

Wang, K. D., Wu, Z., Tufts, L. II, Wieman, C., Salehi, S., & Haber, N. (2024). *Scaffold or crutch? Examining college students' use and views of generative AI tools for STEM education* [Preprint]. arXiv. https://arxiv.org/abs/2412.02653

Proceedings of the Australian Conference on Science and Mathematics Education, The University of Melbourne, 30 September - 2 October 2025, pages 87 - 92, ISSN Number 2653-0481.