INTEGRITY BY DESIGN FOR ETHICAL AI USE IN CDIO ASSESSMENT

Sean M. Cunningham

Department of Mechanical and Automobile Engineering, Technological University of the Shannon, Midwest

David A. Tanner

School of Engineering, University of Limerick

ABSTRACT

Generative artificial intelligence (AI) has created uncertainty in higher education assessment because students can use AI productively for ideation, explanation and technical support, but can also use it to bypass learning. This paper presents the AI Reflection Coefficient Model, an integrity-by-design assessment approach developed within a CDIO-aligned Year 4 Precision Engineering module at Technological University of the Shannon (TUS) Midwest. The model requires students to complete a technical design assignment and submit a structured audio reflection on their use of AI. Reflection quality, assessed using a rubric informed by Gibbs' Reflective Cycle and the Traffic Light System for AI use, acts as a coefficient influencing the final assignment grade. The study used an exploratory practitioner-inquiry design, combining thematic analysis of student reflections with descriptive analysis of rubric scores. Fifteen students participated in the survey and reflection activity. Findings suggest that structured reflection helped make AI use more visible, encouraged students to recognise AI limitations and supported discussion of ethical boundaries. However, students required further scaffolding to move from procedural description towards deeper ethical reasoning. The paper contributes a practical model for embedding reflective academic integrity practices within AI-supported CDIO assessment.

KEYWORDS

Academic integrity, generative AI, CDIO, reflective assessment, engineering education, Standards: 2, 3, 7, 8, 11

INTRODUCTION

Generative artificial intelligence (AI) is now embedded within the academic practices of many students. AI tools can assist with brainstorming, explanation, coding support, literature searching, structuring and design ideation. In engineering education, these affordances are potentially valuable because students are often required to work through complex, open-ended and iterative technical problems. The proportion of students using generative AI tools such as ChatGPT for assessments has jumped from 53% in 2024 to 88% in 2025 (Freeman, 2025). the same tools can also be used to generate text, code, design justifications or assessment outputs in ways that obscure the student's own judgement and learning. This creates a direct challenge for academic integrity because traditional assessment products may no longer provide sufficient evidence of student process, authorship or understanding (Cochrane & Ryan, 2023).

Recent work on AI and academic integrity has emphasised that the problem is not simply technological but educational and cultural. Students may understand academic misconduct in general terms yet still struggle to identify the boundary between legitimate support and inappropriate substitution in practice (Risquez et al., 2013). Moya et al. (2024) similarly highlight the need for proactive institutional and pedagogical responses to AI in higher education. Dawson (2020) frames assessment security as a design issue rather than a matter of surveillance alone, while McCabe (2005) argues that academic integrity is a shared responsibility across the learning community. This perspective is important in AI-supported assessment because detection-based approaches alone are unlikely to cultivate ethical judgment or professional responsibility.

One response to this challenge is to design an assessment that embeds integrity into the learning process. Reflective practice offers a mechanism for making student decision-making visible. Gibbs' Reflective Cycle provides a structured process through which students can describe their actions, evaluate consequences, analyse their choices and plan future practice (Gibbs, 1988; Bulman & Schutz, 1997; McDermott & Anselmo, 2024). In the context of AI, reflection can require students to explain how AI was used, what was accepted or rejected, how outputs were checked, and where human judgment remained essential. This approach shifts attention from simply policing final outputs to examining the learning process through which those outputs were produced.

The CDIO framework is particularly suited to this form of assessment because it emphasises active, applied, and iterative engineering learning across the phases of Conceive, Design, Implement, and Operate. CDIO assessment can foreground process evidence, prototyping, technical decision-making and reflective evaluation. This makes it more resilient to AI misuse than static assessments that focus only on a written report or final answer. AI may legitimately support aspects of a CDIO task, particularly early ideation or clarification, but students must still demonstrate engineering judgement in selecting, implementing and defending their solution.

Existing guidance, such as the AI Assessment Scale and Traffic Light System, helps students distinguish between acceptable and unacceptable AI use (Perkins et al., 2023). Green Zone use may include brainstorming and clarification; Amber Zone use may include structuring or support that requires careful oversight; and Red Zone use includes generating final arguments, final conclusions or complete assignment sections. However, guidance alone does not ensure that students will consistently apply these boundaries. A practical assessment mechanism is required to connect AI-use guidance with student accountability, reflection and grading. This study investigates how a structured AI Reflection Coefficient Model can promote academic integrity, ethical AI use and reflective practice within a CDIO-aligned engineering assessment. The research question is: How can reflective assessment practices promote academic integrity in AI-supported CDIO engineering assessment?

CONTEXT OF THE STUDY

The study was conducted in the Automated Manufacturing Systems module (ENGR0815) at Technological University of the Shannon (TUS) Midwest. The module is delivered to Year 4 Precision Engineering students and includes practical, design-based work aligned with the CDIO framework. Fifteen students participated, representing the full cohort enrolled in the module.

The assessment intervention was embedded in a CDIO-aligned technical design assignment. Students were required to conceive, design, implement, and operate a simple material-handling system or end effector powered by a single NEMA 17 stepper motor and controlled by an Arduino Uno. The system was required to handle a 2 cm cube with precise control within a 30 cm spatial constraint. Deliverables included the design concept, CAD evidence, a physical prototype, the control approach, and a technical explanation of the system.

AI use was permitted for ideation, clarification, coding support, design inspiration and concept exploration. Students were not permitted to use AI to replace final student judgement, final technical argument, final design justification or final submission content. This distinction was communicated using the Traffic Light System (Perkins et al., 2023). Students subsequently submitted structured reflections on their use of AI, primarily in audio format, to explain how AI supported or influenced their assignment process.

AI REFLECTION COEFFICIENT MODEL

The AI Reflection Coefficient Model was developed by the authors for this study as a practical grading mechanism to connect technical performance with reflective evidence of responsible AI use. The model is intended to move academic integrity from a retrospective detection activity into an embedded assessment design feature. Students are assessed not only on the technical quality of their assignment but also on the transparency, criticality and ethical reasoning demonstrated in their AI reflection.

The model combines three elements. First, Gibbs' Reflective Cycle structures the reflection by encouraging students to describe, evaluate and analyse their use of AI and to identify future strategies for responsible use. Second, the Traffic Light System provides a shared vocabulary for distinguishing encouraged, cautionary and prohibited AI practices. Third, a five-criterion rubric assesses the quality of the reflection using the following criteria: Transparency, Critical Reflection, Ethical Awareness, Action Plan for Future Use, and Authenticity. Appendix A, Table 4, outlines the objective grading criteria rubric (Cunningham, S. M., 2025a).

The coefficient calculation was developed by the authors for this study as a practical grading mechanism to connect technical performance with reflective evidence of responsible AI use. The final coefficient grade was calculated as shown in Eqn. (1):

$$[Assignment\ Grade] \times \left(\frac{[AI\ Reflection\ Grade]}{100}\right) = Coefficient\ Grade \tag{1}$$

Example: Students score 80% on the assignment and 65% on the AI reflection, final grade:

$$[80] \times \left(\frac{[65]}{100}\right) = 52\%$$

For example, if a student achieved 80% on the technical assignment and 65% on the AI reflection, the final coefficient grade would be 52%. A minimum threshold was also applied to protect assessment integrity. Where a student scored below 50% on the AI reflection, the student was required to resubmit the reflection or the final report grade was capped at 60%, depending on the assessment context and academic judgement.

The model is not intended to function as an AI detection system. Where discrepancies were suspected between declared AI use and the submitted work, the reflection provided a basis for further dialogue, clarification or reflective review with the student. In this sense, the model functions as a reflective and dialogic integrity tool rather than a mechanism for objectively detecting undeclared AI use.

METHOD

This study used an exploratory practitioner-inquiry design combining qualitative thematic analysis with descriptive quantitative analysis of rubric scores. The intervention was implemented within one final-year engineering module and was designed to examine how structured AI reflection could support ethical engagement with AI in a CDIO assessment context.

Participants were fifteen Year 4 Precision Engineering students enrolled in ENGR0815. All students received guidance on permitted and prohibited AI use through the Traffic Light System and completed the CDIO design assignment described above. Students submitted reflections on AI use after completing the assignment. Reflections were recorded as audio submissions where possible and were subsequently reviewed for evidence of transparency, critical reflection, ethical awareness, future planning and authenticity.

Data sources included student AI reflections, survey responses on AI use, reflection rubric scores and lecturer observations of the assessment process. The survey and reflections were used to identify recurring themes in students' experiences of AI-supported assessment. Thematic coding focused on reported drawbacks of AI, plans for transparent use, initial hesitation, perceived support for learning, effects on critical thinking and reported over-reliance. Reflection rubric scores were analysed descriptively. Six anonymised cases are presented to illustrate variation in student engagement across high, mid-range and low reflection performance. These cases are not used as a statistical sample but as illustrative cases to support the interpretation of the rubric outcomes. Ethical approval was granted by the Faculty of Science and Engineering Research Ethics Committee at the University of Limerick (Approval No. 2024_10_16_S&E). Participation was handled in accordance with the approved ethical protocol, and student examples are anonymised, Cunningham, S. M. (2025a).

Table 1. AI reflection rubric used to assess students' reflective engagement with AI

Criterion	Focus of assessment	Weight
Transparency	Clarity and specificity in documenting how AI was used, with examples and reference to permitted, cautionary or prohibited use.	20%
Critical Reflection	Depth of analysis of AI's impact on learning, including strengths, limitations and student decision-making.	30%
Ethical Awareness	Understanding of originality, academic integrity, responsible AI use and appropriate boundaries.	20%
Action Plan for Future Use	Specificity and practicality of strategies for improving future AI use.	15%
Authenticity	Evidence of a genuine, personal and process-based reflection rather than a formulaic declaration.	15%

RESULTS

The results are reported in two forms. First, cohort-level thematic findings are presented from the full group of fifteen students. Second, six anonymised cases are used to illustrate variation in reflective performance across the rubric criteria. This structure responds to the need to distinguish cohort-level patterns from illustrative examples.

Table 2. Theme frequencies from the full cohort (N = 15)

Theme identified in student reflections/survey	Number of students	Percentage of cohort
Reported drawbacks or limitations of AI	12	80%
Planned transparent AI use in future work	11	73%
Hesitant or unsure at the start of the assignment	7	47%
Reported that AI helped critical thinking	4	27%
Felt that AI-supported learning	3	20%
Reported over-reliance on AI	3	20%

Thematic analysis showed that students recognised both the usefulness and limitations of AI. The most frequent theme was the recognition of AI's drawbacks, identified by 12 students. These drawbacks included inaccurate or shallow outputs, generic suggestions and concerns about over-reliance. Eleven students described plans for a more transparent future AI use, including cross-checking AI outputs, documenting AI support and restricting AI to ideation or clarification. Seven students reported hesitation or uncertainty at the start of the assignment, often because they were unsure where the boundary between support and misconduct lay.

These findings suggest that the reflective task made students more aware of the practical and ethical uncertainties associated with AI use. Students appeared willing to use AI, and many recognised its limitations, but fewer were able to articulate how AI contributed to deeper learning. This indicates the need for stronger scaffolding around reflective and ethical reasoning.

Table 3. Illustrative anonymised AI reflection rubric scores for six students

Student	Transparency (20)	Critical reflection (30)	Ethical awareness (20)	Action plan (15)	Authenticity (15)	Total
Student A	16	22	18	7	13	76%
Student B	9	12	8	7	12	48%
Student C	12	15	8	6	12	53%
Student D	16	21	10	6	13	66%
Student E	18	26	18	10	14	86%
Student F	16	21	12	6	13	68%
Mean	14.5	19.5	12.3	7.0	12.8	66.2%

The illustrative cases show that reflection quality varied considerably. Student E achieved the highest overall reflection grade (86%) and demonstrated strong transparency, critical reflection and ethical awareness. Student A also demonstrated strong reflective engagement but had a weaker future action plan. Students D and F provided honest, technically descriptive reflections but gave less developed attention to ethical reasoning and to future AI-use strategies. Students B and C achieved lower reflection grades because their reflections provided limited evidence of critical evaluation or ethical framing, even though their technical project work may have been stronger.

Across the six illustrative cases, the mean overall AI reflection grade was 66.2%. The strongest criterion was Authenticity, with a mean of 12.8 out of 15 (85.6%), suggesting that students could provide personal accounts of their process. The weakest criterion was Action Plan for Future Use, with a mean of 7.0 out of 15 (46.7%). This suggests that students found it easier to describe what had happened than to translate that experience into a future strategy for responsible AI use. Ethical Awareness also required development, with a mean of 12.3 out of 20 (61.7%). These patterns indicate that reflective AI assessment requires iterative practice and formative feedback if students are to move beyond procedural description.

DISCUSSION

The findings suggest that structured reflection can make AI use more visible in CDIO assessment. Students were required to explain how AI contributed to the Conceive, Design, Implement and Operate phases of their technical work, rather than simply submitting a final product. This process supported academic integrity by creating an assessment space in which students had to account for their use of AI, evaluate its contribution and identify its limitations.

The cohort-level themes show that students were alert to the drawbacks of AI, particularly inaccurate, shallow or generic outputs. This is encouraging because it suggests that students were not simply accepting AI responses uncritically. However, the smaller number of students who explicitly described AI as supporting learning or critical thinking indicates that AI literacy remains uneven. Students may use AI frequently without yet having a mature language to explain how it supports or constrains their learning. This supports the argument that academic integrity guidance must be accompanied by opportunities for practice, feedback and reflective development.

The Traffic Light System provided useful boundary language (Perkins et al., 2023), but the results suggest that it was not yet fully internalised. Several students used AI in ways that were broadly consistent with Green or Amber Zone use, yet did not explicitly refer to the framework in their reflections. This indicates that guidance should not be treated as a one-off briefing. Future iterations should integrate the Traffic Light System into formative checkpoints, peer discussion and rubric feedback so that students learn to apply it as part of their own ethical decision-making.

The coefficient model assigned an assessment value to both the process and the product. This is important in AI-supported assessment because the final artefact alone may not reveal whether students have exercised engineering judgement. By moderating the technical grade through reflective evidence of responsible AI use, the model communicates that how students reach a solution matters. This aligns with CDIO principles because students are expected to justify design decisions, demonstrate implementation choices and evaluate operation in context.

Audio reflections may provide a more personal, process-oriented account of students' decision-making than a written declaration alone. However, it should be understood as a reflective and dialogic integrity tool rather than a detection mechanism. The purpose is not to prove whether AI has been used, but to create a structured opportunity for students to explain their reasoning, acknowledge uncertainty, and discuss ethical boundaries. Where inconsistencies are suspected, the reflection can support further academic dialogue rather than an automatic penalty. The model can also be understood through the lens of Legitimate Peripheral Participation (Lave and Wenger, 1991). Students begin by using AI for relatively low-risk Green Zone activities such as ideation and clarification. Through feedback and reflection, they encounter more complex Amber Zone decisions requiring judgment and accountability. Over time, repeated reflective practice may help students internalise the norms of an academic and professional engineering community. This reinforces McCabe's (2005) view that integrity is a shared responsibility rather than a rule imposed solely on students.

CONCLUSION

This paper presented the AI Reflection Coefficient Model as an integrity-by-design approach for AI-supported CDIO assessment. The model combines Gibbs' Reflective Cycle, the Traffic Light System and a weighted reflection rubric to connect technical performance with evidence of responsible AI use. In the context of a Year 4 Precision Engineering module, the model helped make AI use more visible and encouraged students to consider transparency, ethical awareness and the limits of AI-generated support.

The findings suggest that reflective assessment can support academic integrity by shifting attention from final output alone to the student's learning process, judgment, and accountability. However, the implementation also shows that reflection must be scaffolded. Students recognised AI limitations but often required further support to articulate ethical reasoning and future strategies for responsible use. The weakest rubric area was Action Plan for Future Use, indicating that future iterations should include more structured formative feedback and exemplars. The study is exploratory and limited by its small cohort, single-module implementation and use of illustrative rubric cases rather than full cohort rubric statistics. The findings should therefore be interpreted cautiously. Future work should apply the model across larger cohorts, multiple CDIO modules and repeated assessment cycles, with more robust longitudinal data on AI use, technical performance and reflective development. Despite these limitations, the model offers a practical contribution to CDIO engineering education by embedding academic integrity within the assessment design rather than treating it solely as a post-submission compliance issue.

ACKNOWLEDGEMENTS AND ETHICAL APPROVAL

The authors thank the students in the Automated Manufacturing Systems module at Technological University of the Shannon (TUS) Midwest for their engagement and reflections, which contributed to this study. Ethical approval was granted by the Faculty of Science and Engineering Research Ethics Committee at the University of Limerick (Approval No. 2024_10_16_S&E). The authors also acknowledge the support of academic and administrative colleagues at TUS Midwest and the guidance of the Centre for Transformative Learning at the University of Limerick, particularly Dr Mike Wride. The authors received no financial support for this work.

REFERENCES

Freeman, J. (2025). Student generative AI survey 2025. HEPI Policy Note 61. Higher Education Policy Institute. https://www.hepi.ac.uk/reports/student-generative-ai-survey-2025/

Cochrane, T., & Ryan, T. (2023). ChatGPT and academic integrity: Options for adapting. University of Melbourne Centre for Higher Education. https://melbourne-cshe.unimelb.edu.au/__data/assets/pdf_file/0008/4533218/ChatGPT-and-Academic-Integrity.pdf

Risquez, A., O'Dwyer, M., & Ledwith, A. (2013). Thou shalt not plagiarise: From self-reported views to recognition and avoidance of plagiarism. Assessment & Evaluation in Higher Education, 38(1), 34-43. https://doi.org/10.1080/02602938.2011.596926

Moya, B. A., Eaton, S. E., Pethrick, H., Hayden, K. A., Brennan, R., Wiens, J., & McDermott, B. (2024). Academic Integrity and Artificial Intelligence in Higher Education (HE) Contexts: A Rapid

Scoping Review. Canadian Perspectives on Academic Integrity, 7(3). https://doi.org/10.55016/ojs/cpai.v7i3.78123

Dawson, P. (2020). Defending assessment security in a digital world. London: Routledge. https://doi.org/10.4324/9780429324178

McCabe, D. L. (2005). It takes a village: Academic dishonesty and educational opportunity. Liberal Education, 91(3), 26–31. https://files.eric.ed.gov/fulltext/EJ720381.pdf

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2023). The AI Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. https://doi.org/10.48550/arXiv.2312.07086

Gibbs, G. (1988). Learning by doing: A guide to teaching and learning methods. Oxford: Further Education Unit, Oxford Polytechnic.

Bulman, C., & Schutz, S. (1997). Reflective practice in nursing. Oxford: Wiley-Blackwell.

McDermott, B., & Anselmo, L. (2024). Academic integrity reflection tool: Responsible GenAI use for students in assessments. Taylor Institute for Teaching and Learning, University of Calgary. https://taylorinstitute.ucalgary.ca/resources/Academic-Integrity-Reflection-Tool

Cunningham, S. M. (2025a). AI Reflection Coefficient Model. OSF. https://doi.org/10.17605/OSF.IO/TQUHJ

Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

BIOGRAPHICAL INFORMATION

Dr Sean M. Cunningham is a Lecturer in Mechanical and Automobile Engineering at the Technological University of the Shannon (TUS), Midwest. His research focuses on assessment design, reflective practice and the integration of emerging technologies, particularly artificial intelligence, into engineering education.

Prof David A. Tanner is an Associate Professor in the School of Engineering at the University of Limerick. His work centres on engineering design education, CDIO implementation and assessment innovation in technical curricula.

Corresponding author

Sean M. Cunningham
Technological University of the Shannon
Department of Mechanical and
Automobile Engineering
Moylish, Limerick, Ireland, V94 EC5T
Sean.Cunningham@tus.ie

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivatives 4.0 International License.

APPENDIX A

Table 4: A five-criterion rubric evaluated each student's AI reflection, aligning with the study's LOs and integrity-by-design approach.

Criteria	Excellent (80-100%)	Good (60-79%)	Satisfactory (40-59%)	Poor (0-39%)	(100%)
Transparency	Thorough and clear documentation of AI use, with specific examples aligned to the Traffic Light System.	Describes AI use but lacks some specificity or alignment.	Provide a partial or unclear description of AI use; examples are vague.	Minimal or no description of AI use or misrepresentation of usage.	20%
Critical Reflection	Deep reflection on AI's impact on learning, with thoughtful analysis of strengths and limitations.	Reflects on AI use but lacks depth in the analysis of impacts on learning.	Basic reflection with limited insight into how AI influenced the work.	There is little to no reflection on AI use or overly simplistic responses.	30%
Ethical Awareness	Strong understanding of ethical AI use, including originality and academic integrity.	Some awareness of ethics, but lacks detailed consideration of integrity.	Limited mention of ethical considerations; fails to address key issues.	No engagement with ethics or misuse of AI.	20%
Action Plan for Future Use	A clear and detailed plan for improving AI use in future assignments.	General ideas for future use, but lacking specific strategies.	Basic action plan with limited strategies for improvement.	There is no action plan for future AI use.	15%
Authenticity	Engages deeply with the reflection, providing a genuine and personal response.	There is some engagement, but it lacks depth.	Reflection feels formulaic or lacks personal engagement.	There is little to no personal engagement or insight into AI use.	15%