JUTLP
Journal of University Teaching & Learning Practice

Co-designing for Integrity: A Student Partnership Approach to Developing Ethical AI Literacy Resources in Higher Education

Dr Nantana Taptamat, Marnie Holt, Dom McGrath, and Tiarna McElligott
University of Queensland, Australia

Editors

Section: Special Issues
Editor-in-Chief: Dr Joseph Crawford
Senior Editor: Dr Cassandra Colvin

Publication

Submission: 20 June 2025
Revised: 28 April 2026
Accepted: 29 April 2026
Online First: 9 June 2026

Abstract

Educators and institutions are grappling with how to navigate the integration of generative artificial intelligence (GenAI) while preserving academic integrity and human-centred education. The rapid emergence of GenAI has created a persistent disconnect between institutional policy and practice, leaving students and staff to manage ambiguity in which anxiety often displaces agency. This paper reports on a Design-Based Research (DBR) study that explored how a theoretically informed, co-designed pedagogical strategy can support the development of ethical AI literacy, defined as students' knowledge, reasoning, and regulatory capacities to use GenAI ethically. Drawing on a Student–Staff Partnership (SSP) methodology grounded in regulatory learning theory and emotion regulation, we iteratively developed and formatively evaluated a multimodal educational module within the Faculty of Science of a research-intensive Australian university. Independent quality review provided formative evidence that the co-design process produced resources perceived as authentic, accessible, and empathetically framed. Pilot implementation, however, revealed a critical challenge: an opt-in delivery model had limited reach among the intended audience. From this work we advance four emergent, testable design principles for ethical pedagogy: partner for authenticity, scaffold ethical self-regulation, attend to holistic competence, and employ empathy as pedagogy. The findings suggest a need for educators to embed AI literacy directly into curriculum, moving beyond centralised, opt-in supports.

Practitioner Notes

1. Student anxiety around GenAI can impede ethical reasoning. Pedagogy that attends to wellbeing while building practical AI literacy skills may help students engage with ethical decision-making.
2. Ethical AI literacy is built through scaffolded self-regulation (planning, monitoring, reflecting), not policy memorisation. Authentic, scenario-based activities can scaffold this development by helping students practise planning, monitoring, and reflecting on AI use.
3. Co-design with student partners helps ensure resources target real regulatory challenges and models human-centred educational practice.

Keywords

Generative Artificial Intelligence; Academic Integrity; Co-design; Student-Staff Partnerships; Design-Based Research; Ethical AI Literacy

Citation:

Taptamat, N., Holt, M., McGrath, D. & McElligott, T. (2026). Co-designing for integrity: A theoretically informed approach to developing ethical ai literacy and agency in higher education. *Journal of University Teaching and Learning Practice, 23*(5). https://doi.org/10.53761/m1vv5t24

Introduction

The integration of generative artificial intelligence (GenAI) tools such as ChatGPT, Copilot, and Gemini into higher education presents not only transformative pedagogical opportunities but also ethical and practical challenges (Crawford et al., 2023; Lodge et al., 2024). Students and staff are navigating rapid technological change in a context where institutional regulatory frameworks remain under development (McDonald et al., 2025; Miao & Holmes, 2023). This evolving environment opens space and pressure for educators to design pedagogy that supports learners to use GenAI in ways that enhance, rather than compromise, the integrity and quality of their learning (Bretag, 2016; Floridi & Cowls, 2019).

The challenge is not solely one of rule enforcement. Recent scholarship argues that the most consequential pedagogical question concerns how to foster human agency: supporting learners to make transparent, justifiable decisions about when and how to use AI, to verify outputs, and to exercise oversight aligned with widely accepted AI ethics principles (Adarkwah et al., 2025; Floridi & Cowls, 2019). Such agency requires more than policy literacy; it requires students to develop regulated learning capacities that integrate cognitive judgement with emotional regulation under conditions of uncertainty (Pekrun, 2006; Zimmerman, 2002). Moving beyond ban/allow binaries, this view calls for pedagogy that scaffolds decision-making and disclosure within authentic learning and assessment tasks (Crawford et al., 2023; Lodge et al., 2024; Rudolph et al., 2024).

Empirical evidence at our institution, a large research-intensive Australian university, indicates the pedagogical significance of this challenge. Qualitative data gathered during the project through focus groups with science students revealed uncertainty that inhibited agency: students described fear of unintentional academic integrity breaches due to inconsistent guidance, reliance on informal support networks such as peers and social media, and avoidance behaviours that limited productive engagement with AI tools. These patterns are consistent with broader student-experience research documenting both enthusiasm about GenAI as well as significant concern about reliability, fairness, and integrity (Chan & Hu, 2023; Henderson et al., 2025). The pedagogical question is therefore how educators can respond constructively to these conditions.

Despite a growing body of scholarly commentary on GenAI in higher education, there remains a paucity of theoretically informed, empirically documented pedagogical design cases that show how educators can translate ethical principles into accessible, student-validated learning resources. Specifically, little is known about how student–staff partnership can be operationalised to produce ethical AI literacy materials that respond to the emotional as well as cognitive demands students experience. This study addresses this gap.

> **Research Question:** How can Student–Staff Partnership support the design, formative evaluation, and implementation planning of ethical AI literacy resources in higher education?

Specifically, this study asks:

> **RQ1:** What student needs and regulatory challenges informed the design of the ethical AI literacy resource?

RQ2: How did student–staff co-design shape the content, tone, and pedagogical structure of the resource?

RQ3: What formative evaluation and implementation insights emerged from the pilot and review cycles?

To attend to these research questions, we initiated the development of the Generative AI Essential Guide: Unlocking Learning in the Sciences (hereafter, the FoS GenAI Guide), a multimodal educational module co-designed with students. A practitioner-facing case study of this initiative has been published elsewhere (Taptamat et al., 2025); the present article extends that account by providing a theoretically framed Design-Based Research analysis of the co-design process, formative evaluation, and implementation insights. The paper is structured across four sections. The Literature section synthesises scholarship on ethical tensions in GenAI, co-design and student–staff partnership, regulated learning, emotion regulation, and design-based research. The Method section details the Design-Based Research approach, the five DBR cycles, participant safeguards, data sources, and the reflexive thematic analysis used to interpret qualitative data. The Results section is organised around the three sub-questions: student needs and regulatory challenges, student–staff co-design contributions, and formative evaluation and implementation insights. The Discussion considers emergent design principles, practical and theoretical implications, and the scope of claims made by the study.

The significance of this research for learning and teaching practice is threefold. First, it offers a worked example of how educators can develop ethical AI literacy resources grounded in both student experience and regulatory learning theory, providing a model that may be adapted across disciplinary contexts. Second, it identifies a set of emergent design principles that can guide others developing similar resources and that are positioned for empirical validation in subsequent research. Third, it surfaces an implementation paradox: quality alone is insufficient to ensure reach. This suggests that AI literacy may need to be embedded within the curriculum rather than relying solely on centralised, opt-in supports. Taken together, the study contributes a human-centred pedagogical response to one of the most pressing practical challenges facing university teachers.

Literature

Generative AI, integrity, and the pedagogical turn

The ethical implications of GenAI in higher education centre on academic integrity, authenticity, bias, and fairness (Rudolph et al., 2024). GenAI tools can scaffold learning through low-stakes feedback, ideation, and formative dialogue (Rana et al., 2025; Smith et al., 2025), yet they also risk undermining originality, amplifying inequities, and displacing effortful cognitive work necessary for deep learning (Lodge et al., 2024). Institutional responses have often prioritised detection, surveillance, and punitive framings (Tillmanns et al., 2025). Recent syntheses indicate that policy guidance remains incomplete and inconsistent, with definitions of acceptable use varying across institutions and within the same institution across courses (Gonsalves, 2025; McDonald et al., 2025). Students and staff therefore experience inconsistency, ambiguity, and uncertainty in their daily practice (Chan & Hu, 2023).

Scholars increasingly argue for a pedagogical turn in which integrity is reframed from compliance to empowerment (Bretag, 2016; Crawford et al., 2023; Lodge et al., 2024). Ethical

AI literacy, in this framing, is not the memorisation of rules but the capacity to make principled decisions under uncertainty. This includes recognising when AI enhances versus undermines learning; evaluating outputs critically for accuracy, bias, and relevance; developing strategies for transparent attribution; and enacting these principles with confidence in authentic academic tasks. For this study, we therefore conceptualise ethical AI literacy as comprising three mutually reinforcing dimensions: (a) knowledge of how integrity principles apply to generative AI; (b) the capacity for ethical reasoning in ambiguous contexts; and (c) the agency and confidence to enact these principles in real tasks (Bretag, 2016; Eaton, 2023; Lodge et al., 2024). This conception places cognitive, metacognitive, and affective capacities on equal footing and anticipates the regulatory framing developed below.

Co-design and Student–Staff Partnership

Co-design with students is a well-established pedagogical response to the challenge of developing relevant, legitimate, and accessible educational resources (Bovill, 2020; Healey et al., 2014; Matthews, 2017). Student–Staff Partnership (SSP) is one influential articulation of co-design in higher education. Matthews (2017) advances five propositions for genuine partnership practice: fostering inclusive partnerships, nurturing power-sharing relationships, accepting partnership as a process with uncertain outcomes, engaging in ethics of reciprocity, and honouring the transformative potential of partnership. Healey and others (2014) similarly emphasise that partnership is not a pedagogical technique but an ethic of engagement that challenges traditional hierarchies between staff and students.

Where students are positioned as genuine partners rather than passive recipients or sources of data, they bring lived experience that surfaces blind spots in staff-led design, co-create scenarios that reflect authentic dilemmas, and cultivate a sense of ownership that is itself pedagogically important (Bovill, 2020). In the context of GenAI, co-design allows students to articulate their concerns, critique proposed guidelines, and contribute to the production of authentic examples grounded in their disciplinary contexts (Rahimi, 2025; Rana et al., 2025). Such dialogic practices are relevant to critical thinking and communication capacities. Importantly, the co-design of ethical AI literacy resources also models the human-centred, collaborative pedagogy needed to balance automation with human judgement ethically, thereby enacting, at the level of process, the ethical stance the resources seek to cultivate.

Regulatory learning as a theoretical framework

Ethical AI use in academic contexts fundamentally requires regulated learning: the self-generated thoughts, feelings, and behaviours oriented towards attaining goals (Zimmerman, 2002, p. 65). We draw on three interconnected modes of regulation. Self-regulated learning (SRL) comprises cyclical phases of forethought (planning, goal-setting), performance (monitoring, strategy use), and self-reflection (evaluation, adaptation) (Panadero, 2017; Zimmerman, 2002). Co-regulated learning (CoRL) refers to transitional processes in which a more knowledgeable other temporarily provides regulatory support that learners gradually internalise (Hadwin et al., 2011). Socially shared regulated learning (SSRL) involves groups jointly regulating through negotiated planning, coordinated monitoring, and collective reflection (Järvelä & Hadwin, 2013).

These modes map directly onto the pedagogical challenges identified in our needs analysis. Students report difficulties in forethought (deciding whether and how to use GenAI), performance (monitoring outputs and staying aligned with learning goals), and self-reflection (evaluating and disclosing AI use). They also rely heavily on informal, often unreliable co-regulators (peers, social media), suggesting a maladaptive co-regulation pattern that well-designed resources might productively redirect. Empirical research has shown that learners typically require explicit scaffolds to transfer SRL strategies to digitally mediated tasks (Broadbent et al., 2020; Broadbent & Poon, 2015; Panadero, 2017). Ethical AI literacy resources must therefore do more than provide rules: they must scaffold the regulatory phases themselves.

Emotion regulation and the affective dimension of ethical reasoning

Our needs analysis identified pervasive anxiety as a defining feature of students' experience of GenAI. This finding points to the importance of emotion regulation within the ethical AI literacy framework. Pekrun's (2006) control–value theory of achievement emotions provides a useful lens: anxiety arises when students perceive low control over outcomes that are valued (for example, academic integrity, fair grading), and is associated with reduced task engagement, narrower attention, and impaired working memory. Artino and Jones (2012) similarly show that negative activating emotions such as anxiety tend to suppress the deep cognitive and metacognitive strategies that SRL relies on, while positive emotions support them. Mega and colleagues (2014) demonstrate empirically that emotion and motivation mediate the relationship between SRL strategies and academic achievement.

For the design of ethical AI literacy resources, this evidence has a clear implication: cognitive scaffolds for ethical reasoning are unlikely to be effective if the affective conditions for engagement are absent. Resources must pair cognitive scaffolds (e.g., worked examples, checklists, reflection prompts) with affective ones (e.g., empathetic tone, recognition of uncertainty, normalisation of the challenge) to create the psychological safety required for learners to engage productively with the cognitive demands of ethical reasoning.

Design-Based Research as methodological grounding

Design-Based Research (DBR) is suited to the iterative development and refinement of educational interventions in authentic settings, while simultaneously advancing pedagogical theory (McKenney & Reeves, 2018; Wang & Hannafin, 2005). DBR is characterised by (a) a focus on authentic problems of practice; (b) collaboration between researchers and practitioners; (c) iterative cycles of analysis, design, development, and implementation; and (d) the production of design principles that are testable in subsequent research. These commitments shaped the five DBR cycles that produced and refined the FoS GenAI Guide, described in the Method section that follows.

Method

The study used Design-Based Research (DBR) as the overarching methodological framework. DBR was selected because the research aim was both practical (to produce a useful educational resource) and theoretical (to advance design principles that can guide comparable work in other settings), and because DBR explicitly accommodates the iterative, collaborative nature of the design process (McKenney & Reeves, 2018; Wang & Hannafin, 2005). The five DBR cycles were guided by evidence about where student regulation was breaking down; each cycle targeted those phases with progressively refined scaffolds, with evaluation data informing subsequent iterations. Table 1 maps the regulatory framework to the Guide components that emerged from this process.

Table 1 Mapping the Regulatory Framework to FoS GenAI Guide Components

Regulatory target (SRL phase)	Student regulatory challenge	Intervention component (artefact)
Forethought (plan and task-fit)	Deciding whether/when/how to use GenAI for a specific task	Module 1 + task-fit scenarios
Performance (monitor and verify)	Checking accuracy, bias, alignment with learning goals	Module 2 verification checklist + exemplars
Performance (stay on learning)	Avoiding drift towards over-reliance or task substitution	Scenario guides examining ethics, risks, and impacts on learning
Self-reflection (evaluate use)	Judging whether AI use was effective and ethical	Reflection prompts
Transparency (disclose and attribute)	Documenting AI assistance responsibly	Disclosure template + examples
Emotion regulation (enabler across phases)	Managing anxiety about misconduct	Empathy-framed guidance throughout

The research took place in the Faculty of Science at a research-intensive Australian university between November 2023 and August 2025. Ethical approval was obtained from the University's Human Research Ethics Committee (2024/HE001346), with voluntary, written informed consent obtained from all participants; data were de-identified prior to analysis, and student partners were remunerated through a university programme. The development of the GenAI educational module was prompted by student feedback indicating pervasive anxiety about appropriate AI use and followed five iterative Design-Based Research (DBR) cycles. Cycle 1 (November 2023–April 2024) involved an analysis phase drawing on a 2023 Science Faculty survey of 322 science students, a desktop policy analysis, and staff roundtables. Cycle 2 (April–June 2024) convened the first Student-Staff Partnership (SSP) team, involving four student partners (two undergraduate, two postgraduate) and three staff who worked 50 hours over 15 weeks to co-create Module 3 content and provide feedback on usability and tone. In Cycle 3 (May–July 2024), five students trained in the Student-Led Observation for Course Improvement (SLOCI) programme and 11 staff reviewers evaluated the draft Guide using a TELAS-informed framework. Cycle 4 (July–October 2024) piloted the Guide across four courses (N = 235); this yielded 12 survey responses and engaged 17 students in faculty-wide focus groups to examine reach and perceived value. Notably, none of the focus-group participants had previously accessed the Guide. Finally, Cycle 5 (February–July 2025) convened a second SSP team of four student partners (three returning; one undergraduate, three postgraduate) and three staff to co-design five additional practical scenarios over 15 weeks. Staff consultants, including course coordinators, learning designers, and librarians provided ongoing guidance across all cycles. The comparatively small pilot and survey samples reflect the opt-in delivery model whose limitations the study interrogates, and the cycles, participants, and outcomes are summarised in Table 2.

Table 2

DBR Cycles for the Development of the Generative AI Essential Guide

Cycle	Timeline	Participants	Focus/ Outcomes
1. Analysis and Initial Development	Nov 2023–Apr 2024	Faculty staff	Identified policy–practice gaps; established priorities
2. Student–Staff Partnership and Refinement	Apr–Jun 2024	SSP team (n = 7)	Co-created content; shaped tone and usability
3. Comprehensive Review and Enhancement	May–Jul 2024	Student review group (n = 5); faculty staff	Refined clarity, accessibility, and policy alignment
4. Implementation and Evaluation	Jul–Oct 2024	Students (N = 235); focus groups (N = 17)	Perceived usefulness among engaged students; established ongoing need for scaffolding of ethical self-regulation
5. Expansion and Scenario Development	Feb–Aug 2025	SSP team (n = 7); academic reviewers	Revised the Guide; added five scenarios; academic review; planning faculty rollout

The intervention consisted of an online Generative AI Essential Guide for science students. The Guide comprised three modules: Module 1 (Foundation: capabilities and limitations of GenAI), Module 2 (Responsible Use: navigating ambiguity and making principled decisions), and Module 3 (Practical Application: ethical use in authentic science tasks), interleaving short activities, scenarios, and knowledge checks. Scenario guides followed a consistent pattern (purpose; when to use / when to avoid; risks and ethics; learning impact; recommended practice) explicitly mapped to forethought, performance monitoring, and self-reflection. The structure of the guide is summarised in Table 3.

Data collection and analysis

The study drew on multiple evidence sources. The post-module pilot survey elicited Likert-scale and open-ended responses on perceived value and clarity of the Guide. Focus-group transcripts captured students' experiences of, and uncertainties around, GenAI use; focus groups were conducted by SLOCI students using a semi-structured interview guide and lasted approximately 45–60 minutes. Student partners produced written reflections and design artefacts (draft content, annotated feedback, meeting notes) throughout the partnership cycles. SLOCI reviewer reports provided formative evaluation of usability, clarity, and tone. Staff consultation notes documented expert and practitioner feedback across cycles. Triangulation across these qualitative sources was used to strengthen the credibility of the design-process analysis (Lincoln & Guba, 1985).

Table 3

Structure of FoS GenAI Guide

Module	Focus	Key Content and Activities	Pedagogical Rationale (Informed by Co-Design)
1. Foundation	Understanding capabilities and limitations	Definitions; types of GenAI tools; capabilities and constraints; prompt engineering basics	Build foundational knowledge; establish realistic expectations about what GenAI can and cannot do
2. Responsible use	Navigating ambiguity and making principled decisions	Ethical principles; institutional integrity policies; evaluation strategies for AI outputs; referencing guidelines	Provide clarity on rules; establish shared understanding of responsible use; address forethought planning needs
3. Practical Application	Applying AI ethically in specific contexts	Discipline-specific examples (e.g., “programming assignment guide”); hands-on critical evaluation activities	Scaffold CoRL through guided practice; address anxiety with empathetic tone; ensure authenticity through student-generated scenarios; support all three SRL phases

Focus-group transcripts and open-ended survey responses were analysed using reflexive thematic analysis (Braun & Clarke, 2006, 2019). Reflexive thematic analysis was selected because the study sought to generate interpretive insight into students’ experiences of GenAI rather than to produce a taxonomy of fixed codes. Consistent with this paradigm, we did not pursue inter-rater reliability, which rests on epistemological assumptions incompatible with reflexive TA (Braun & Clarke, 2019). Instead, we established trustworthiness through the procedures recommended for qualitative work of this kind (Lincoln & Guba, 1985; Nowell et al., 2017): reflexive journalling by analysts; memoing throughout coding to trace the evolution of interpretations; peer debriefing within the research team; and an audit trail linking raw data to coded extracts, candidate themes, and final themes. Two SLOCI student analysts led the initial coding. Both had completed SLOCI training in qualitative and quantitative methods, and were further briefed on reflexive TA and the specific research question before commencing analysis. The first author supervised the analytic process, reviewing raw data, candidate themes, and the final analysis to ensure interpretive consistency with the dataset as a whole. Likert-scale survey items were summarised descriptively; the small survey sample precludes inferential analysis, and descriptive summary is reported for transparency rather than as an effectiveness claim.

Results

Results are organised around the three sub-questions: the student needs and regulatory challenges that established the design context, the student–staff co-design contributions that translated student input into pedagogical decisions, and the formative evaluation and implementation insights that tested design quality and surfaced an implementation paradox.

Student needs and regulatory challenges informing design, addressing RQ1

The initial draft of the Guide was informed by a 2023 survey of 322 science students (FoS SLOCI, 2023), which revealed widespread GenAI use (74.5%), ethical concern about plagiarism and reliability, and a clear demand for institutional guidance. Roundtable discussions with faculty staff validated these findings. Three design commitments followed: a student-centred approach to address practical and ethical concerns while enhancing understanding; stakeholder collaboration to integrate diverse perspectives; and multimodal content delivery to support engagement and reflective practice (Mayer, 2020). Focus-group analysis across Cycle 4 identified three interlocking regulatory challenges that shaped the design.

First, students described pervasive anxiety about inadvertent academic integrity breaches. This anxiety was not driven by an intent to cheat but by unclear institutional and course-level boundaries. One student reported, "I'm scared of getting academic penalties from [Turnitin] where they could detect AI" (FG2, Student 2). Another explained the behavioural consequence: "currently I'm scared about getting in trouble for using it, so I just don't use it" (FG3, Student 1). Such avoidance represents a disruption of SRL forethought, consistent with the emotion–regulation literature suggesting that anxiety narrows cognitive engagement and suppresses strategy use (Artino & Jones, 2012; Pekrun, 2006).

Second, students reported significant inconsistency in GenAI guidance across courses, which made it difficult to engage in forethought planning. Beyond integrity concerns, students raised fairness and equity issues. Some described feeling disadvantaged by peers with greater digital literacy or access to paid AI tools: "It's unfair because they get a higher score than people who don't use AI. Then what's the point of going to school and studying?" (FG2, Student 2). Others noted bias in AI outputs, especially for non-English queries: "When I type the same question in Chinese, it says it doesn't know the answer, but in English, it gives the correct answer" (FG1, Student 2). These concerns connect the individual regulatory challenge to broader structural conditions shaping how students can and cannot engage with GenAI ethically.

Third, students relied heavily on informal, often unreliable sources of co-regulation. Peers, friends, and social media platforms such as TikTok were named as primary sources of learning about GenAI use: "Mostly from friends who have more understanding of AI" (FG1, Student 1); "TikTok helped me, gave guidance" (FG2, Student 3). Notably, none of the 17 focus-group participants had encountered the Guide itself prior to the focus group: "No, this is the first time I've ever [heard about it]" (FG5, Student 2); "First time seeing this module" (FG1, Student 3). This finding pointed to a significant discoverability problem with the centralised, opt-in delivery model. Students also explicitly requested course-embedded guidance rather than separate resources, and observed that the complexity of AI declaration requirements deterred use: "I found that it seems like a lot of effort. So, I sort of purposely haven't used AI because of that" (FG3, Student 1). Taken together, the needs analysis surfaced a layered regulatory problem, including affective (anxiety), cognitive (ambiguity), and social (maladaptive co-regulation) that the design needed to address.

Student–staff co-design contributions, addressing RQ2

Student–Staff Partnership translated the needs analysis into a responsive educational design. Three features of the co-design process illustrate how student input critically shaped the resource. Initial staff-led drafts were text-heavy and conceptually oriented, framed primarily around definitions and rules. Student partners argued that this framing would not engage their peers and would fail to address the affective dimension of the student experience. They advocated for a substantive redesign centred on interactive scenarios, short videos, and concrete examples grounded in disciplinary practice. The co-creation of ethical scenarios, drawn from authentic student dilemmas such as whether and how to use GenAI for exam preparation, became the central pedagogical strategy for scaffolding self-regulation.

A representative comparative example illustrates the shift. The initial staff-authored framing for an exam-preparation task quoted institutional rules: “Students must not use generative AI to produce answers to assessment tasks where AI use is not permitted. Doing so may constitute academic misconduct under the University’s policy.” After co-design, the equivalent content was reworked by student partners into a scenario-based format. The revised version opened with a short narrative of a student considering GenAI for revision; presented a decision table with “when to use” and “when to avoid” prompts; included an ethics-and-risk reflection; and concluded with recommended practice framed in supportive language (for example, “if you’re unsure, check your course profile and ask your course coordinator — it’s always better to ask”). The revision retained the integrity principle but reframed it from a prohibition into a decision-making aid. This change illustrated a broader shift in tone, from compliance to empowerment, that student partners consistently advocated across all modules.

Student partners also shaped tone and accessibility. Recognising the anxiety identified in Phase 1, they argued for a non-judgemental, empathetic tone that acknowledged student fears rather than focusing solely on compliance. They championed design choices that reduced cognitive load; one tangible outcome was the development of “a concise, readable table format for levels of GenAI use” (Student Partner 1 reflection). They ensured authenticity by drawing on lived experience: Module 3 (Practical Applications) integrated examples relevant to science students, including coding support, data analysis, and literature review, providing context-specific guidance that moved beyond generic advice. As one student partner reflected, “Being part of the design helped me see that integrity isn’t just about rules. It’s about making good choices. And we needed the Guide to show that in a clear and supportive way.”

Formative evaluation and implementation insights, addressing RQ3

Evaluation occurred in two strands: independent formative quality review and pilot implementation. The draft Guide underwent formative evaluation by the SLOCI student review group (N = 5) using a TELAS-informed framework (ASCILITE TELAS, n.d.), with four professional and seven academic staff providing additional expert feedback. This review provided formative evidence that the co-design process had produced resources perceived as accessible, relevant, and empathetically framed. Reviewers strongly endorsed the multimodal design, praised the context-specific guidance, and highlighted the empathetic, non-judgemental tone. One reviewer noted that the Guide “is quick, informative, and doesn’t demonise or blame students” (SLOCI Reviewer 3). Another commented that it “shows how it can be used to help learning in an ethical manner” (SLOCI Reviewer 4). Reviewers reported that the “non-judgemental” language lowered perceived anxiety when engaging with the material.

Reviewers also identified areas for improvement. They recommended streamlining Module 1 and reducing redundancy, which led to a 20% condensation of introductory content, and suggested additional interactive visuals and downloadable resources (checklists, templates), which were subsequently developed. One reviewer articulated a residual concern that mirrored the anxiety documented in Phase 1: “I understood general principles but if it came to writing an assignment with AI, I’d still feel quite nervous about being penalised accidentally” (SLOCI Reviewer 4). In response, specific worked examples of acceptable and unacceptable AI use were added. This feedback nevertheless highlighted a constraint on what module design can achieve: module designers cannot fully resolve policy ambiguity that exists at institutional and course levels. Crucially, because SLOCI reviewers were not involved in development, their endorsement of the features advocated by student partners, including empathetic tone, multimodal presentation, scenario-based guidance, and the integration of ethical reasoning with task support, provides independent formative support for the value of the co-design process.

Pilot implementation, by contrast, surfaced a critical problem of reach. The Guide was made available to Faculty of Science students in four courses (N = 235) through the learning management system in Semester 2, 2024, and course coordinators were informed of its availability and encouraged to mention it as a supporting resource. Direct engagement was limited: the post-module survey yielded 12 self-selected responses, 10 of the 12 respondents agreed that the module was valuable. While these responses suggest that the co-design approach produced valued content among those who engaged, the sample size is too small to support any claims about effectiveness, acceptability, or reach; these descriptive data are reported for transparency rather than as impact evidence.

The more telling finding emerged from the concurrent focus groups. Seventeen students recruited faculty-wide, including students enrolled in pilot courses and those not enrolled, were asked whether they had encountered the Guide. None had. This gap between design quality and student reach indicates a limited reach of the centralised, opt-in delivery model. Students consistently requested course-embedded guidance; the value of a resource is arguable if students never encounter it. The implementation challenge reflects an assumption common in institutional responses to GenAI that centralised, opt-in literacy supports are sufficient. Our findings suggest that this assumption is incorrect and that responsibility cannot rest solely with students. Teachers may need to integrate resources into course design at the point of need.

Discussion

The study explored how Student–Staff Partnership can support the design, formative evaluation, and implementation planning of ethical AI literacy resources in higher education. In addressing this question, the work makes three interrelated contributions that are worth considering together. First, it offers an empirical specification of students’ self-regulatory challenges when attempting to use GenAI ethically, moving beyond general observations that “students are confused” to identify targetable design needs: disrupted forethought under affective load, limited performance monitoring in the absence of institutional scaffolds, and constrained self-reflection in the face of complex disclosure requirements. These patterns align with research showing that learners often require explicit supports to transfer SRL strategies to digitally mediated tasks (Broadbent & Poon, 2015; Panadero, 2017), and with research on the suppressive effect of anxiety on metacognitive strategy use (Artino & Jones, 2012; Mega et al., 2014; Pekrun, 2006). Second, it documents how Student–Staff Partnership can transform theoretically sound draft resources into authentic, accessible, empathy-centred learning tools; independent formative review provided evidence that the process achieved its intended goals (Bovill, 2020; Healey et al., 2014; Matthews, 2017). Third, it surfaces an implementation paradox: well-designed, student-validated resources may not achieve their intended pedagogical impact unless they are integrated into the curriculum. This paradox has significant consequences for how institutions respond to GenAI, which we discuss below.

The implementation paradox is usefully interpreted through implementation science literature. Implementation science, which studies the conditions under which evidence-based practices are taken up and sustained in real-world settings, has long cautioned against the assumption that quality of design is sufficient for adoption (Damschroder et al., 2009). Fidelity frameworks distinguish between the quality of an intervention and the conditions under which it is delivered, noting that reach (the proportion of the target population exposed) and adoption (by implementers) are typically more sensitive to contextual factors such as visibility, workflow integration, perceived relevance, and institutional support than to intrinsic intervention quality. Our findings make this distinction clear. Independent reviewers provided formative evidence of the Guide's quality; the pilot exposed the weakness of the delivery mechanism. Opt-in dissemination assumes a motivated, resourced student seeking out support, which is the opposite of the student most in need of ethical AI literacy guidance, including the one who is anxious, avoidant, and overloaded. Relying on such a mechanism places the burden of ethical navigation on students, and is therefore not only ineffective but arguably inequitable.

A second interpretive lens strengthens this conclusion. If anxiety suppresses SRL forethought, then students most affected by affective load are precisely those least likely to undertake the deliberate, self-initiated search for help that opt-in resources require (Artino & Jones, 2012; Pekrun, 2006). This is a structural, not individual, problem. Embedding ethical AI literacy within course design at the point of need, including as part of assessment briefs, tutorial activities, or task-specific prompts, relieves students of the discovery burden and brings support into the contexts where regulation is actually taking place. We advance this as a plausible hypothesis requiring empirical test; whether mandated embedding improves outcomes, engenders resentment, or produces superficial compliance is an open question.

It is important to be explicit about what the present study does and does not measure. The study measures the quality of the design process and its outputs as judged by independent reviewers, the relevance and authenticity of the resource as judged by student partners, and the reach of an opt-in delivery model. It does not measure student learning outcomes, changes in ethical reasoning or self-regulatory capacity, or behavioural change in authentic academic tasks. The 12 pilot survey responses indicate perceived usefulness among a self-selected, engaged subsample, and cannot be generalised. The study therefore contributes a worked design case, a set of emergent design principles, and a diagnosis of an implementation problem, rather than an impact evaluation. Effectiveness claims require a subsequent phase of research in which the Guide is embedded in coursework and outcomes are measured against comparison conditions.

Emergent design principles

In DBR, design principles are best understood as hypotheses derived from iterative refinement, positioned for subsequent empirical evaluation (McKenney & Reeves, 2018). Four principles emerged from our design process and formative evaluation. They are presented here as testable propositions rather than validated prescriptions, and each is paired with the kind of empirical work that would be required to test it.

The first principle, partner for authenticity and agency, holds that co-design with students is essential for developing ethical AI literacy resources that reach and engage their intended audience. The Student–Staff Partnership in this study reshaped both content (from theory to scenarios) and tone (from compliance to empathy), and independent reviewers identified these features as strengths. The SSP process also modelled the human-centred, collaborative pedagogy the resources themselves sought to cultivate, aligning process with product. The testable claim is that SSP-developed resources will show greater perceived authenticity, engagement, and use than comparable staff-only resources; this is amenable to comparative empirical test.

The second principle, scaffold ethical self-regulation through scenarios, holds that contextualised scenarios are more effective than abstract rule presentations for supporting learners to rehearse forethought, performance monitoring, and self-reflection in GenAI-rich tasks. The co-created scenarios explicitly map to the phases of SRL and provide the targeted support that research indicates is often required to transfer SRL skills to digital environments (Broadbent et al., 2020). The testable claim is that scenario-based scaffolds will produce stronger regulatory engagement and better-justified AI-use decisions than rule-based instruction.

The third principle, attend to holistic competence, holds that ethical AI literacy cannot be separated from student wellbeing. Pervasive anxiety indicates that the affective conditions for engagement must be designed for explicitly (Artino & Jones, 2012; Mega et al., 2014; Pekrun, 2006). The testable claim is that resources combining cognitive scaffolds with affective framing (empathetic tone, normalisation of uncertainty) will show greater engagement and learning gain than resources that address cognition alone.

The fourth principle, employ empathy as pedagogy, holds that an empathetic, non-judgemental tone is a pedagogical choice. In an environment characterised by fear, punitive framings are counterproductive because they compound the affective load already suppressing engagement. Empathy as pedagogy builds the trust required for students to enact ethical agency. The testable claim is that empathetically framed resources will show higher uptake and more productive engagement with difficult content than identically structured resources delivered in a compliance-focused register.

Practical implications

Several practical implications follow for educators and institutions, framed as hypotheses that warrant validation in implementation research. First, AI literacy supports are more likely to reach learners when they are embedded within courses and assessments than when provided as centralised, opt-in resources. Course-embedded scaffolds meet learners at the point of need and reduce the discoverability barrier that our pilot found to be decisive.

Second, discipline-specific scenarios and templates make authorised and transparent use of GenAI visible in context, and are more likely to be used than generic guidance. Third, student representatives should be partners in both design and dissemination of AI literacy resources; authenticity is necessary but not sufficient for reach.

Fourth, institutional policy ambiguity heightens the responsibility of individual educators to provide clear, course-specific expectations for authorised use and disclosure; this is an immediate practical step educators can take while waiting for policy clarity. Each of these implications is best understood as a hypothesis requiring empirical test rather than a validated prescription.

Theoretical implications

The study also contributes to theoretical debates in three ways. First, it extends the application of regulated learning theory to ethical AI use, specifying how the phases of SRL map onto the ethical reasoning required of students using GenAI. This application foregrounds the cyclical, learnable nature of ethical AI use and reframes compliance-oriented pedagogy as a pedagogy of regulation.

Second, it strengthens the theoretical case for integrating emotion regulation with cognitive SRL in digital learning design. The empirical pattern we observed, including anxiety disrupting forethought and reflection, supports Pekrun's (2006) and Artino and Jones's (2012) contention that affective and cognitive regulation should not be treated as separable. Third, it illustrates how Student–Staff Partnership can function not only as a design methodology but as a theoretical enactment of human-centred pedagogy: the process through which a resource is produced can itself model the ethical stance the resource seeks to cultivate. This aligns with Matthews's (2017) proposition that partnership is an ethic of engagement, not merely a technique, and offers a worked example of that claim in the context of GenAI.

Several constraints limit the claims this study can make. The study provides evidence of the design process and formative quality rather than impact evidence; it does not demonstrate that the Guide improves students' ethical reasoning, self-regulatory capacity, or integrity practices in authentic academic tasks. The pilot survey sample was small and self-selected, and cannot support claims about acceptability, effectiveness, or broader reach. Development occurred within a particular disciplinary context (science), institutional culture, and regulatory environment (Australian higher education); student partners brought science-specific perspectives, the scenarios reflect science tasks, and SLOCI reviewers evaluated from a science student standpoint. Transfer to humanities, professional programmes, or different national and regulatory contexts requires validation.

Co-design carries its own set of methodological risks that warrant explicit reflection. Student partners were self-selected through an open call emphasising interest in educational design and AI ethics, and are unlikely to be representative of the full student population, in particular of students most disengaged or anxious about GenAI. Sustained engagement in partnership may also align partners' perspectives more closely with staff priorities over time, potentially dampening the disruptive critique that makes partnership valuable (Matthews, 2017). We sought to mitigate these risks by actively seeking dissenting perspectives within partnership meetings, using anonymous SLOCI reviewers (who were not partners) for independent evaluation, and triangulating partner reflections with focus-group data from students who had no relationship to the design process. We acknowledge, however, that a degree of design bias is inherent to any co-design study and that this should be weighed when interpreting our findings. Finally, the study hypothesises that course-embedded integration would improve reach, but does not test this claim; whether mandated completion improves outcomes, engenders resentment, or creates superficial compliance requires implementation research comparing embedded and opt-in conditions across disciplinary settings.

Conclusion

Generative AI presents complex ethical challenges that students often experience as anxiety, placing educators in a difficult pedagogical position. Framing ethical AI literacy as a regulated learning construct that integrates cognitive, metacognitive, and affective capacities clarifies how pedagogy can move beyond compliance towards the development of ethical judgement. Through an iterative DBR methodology centred on Student–Staff Partnership, we developed a Generative AI Essential Guide that operationalises this framework, and independent reviewers judged the resulting module authentic, usable, and empathetically framed. Four emergent, testable design principles follow from this work: partner for authenticity and agency; scaffold ethical self-regulation; attend to holistic competence; and employ empathy as pedagogy. While the design quality was formatively validated, the pilot exposed a critical implementation paradox: the opt-in approach used in this study had limited reach, and the students most affected by anxiety are the least likely to find them. This highlights an ethical imperative for educators and institutions to adopt proactive, integrated pedagogical responses. By embedding ethical AI literacy within the curriculum, and by partnering with students in both design and dissemination, institutions can support educators to navigate one of the most pressing ethical landscapes in contemporary higher education.

Acknowledgements

The authors disclose that they have no actual or perceived conflicts of interest and that they have not received any funding for this manuscript beyond resourcing for academic time at their university. Each student partner received $1,000 for their involvement through the Student–Staff Partnership Projects. ChatGPT and Claude were used only for surface-level language editing (grammar and clarity). The intellectual property remains that of the authors. No concepts, analyses, data, or references were generated by AI. All text was reviewed and approved by the authors, and no confidential or identifiable data were entered into the tool. The authors have not used artificial intelligence for ideation, study design, data analysis, or substantive writing, consistent with Crawford et al. (2023). We thank the student partners, focus-group participants, and the academic and professional staff who contributed feedback and helped shape the Guide throughout the project. The authors declare the following CReDIT statement: Conceptualisation: Nantana Taptamat, Dom McGrath, Marnie Holt, Tiarna McElligott; Methodology: Nantana Taptamat, Marnie Holt; Formal analysis: Nantana Taptamat; Investigation: Nantana Taptamat, Marnie Holt, Tiarna McElligott; Writing – original draft: Nantana Taptamat; Writing – review & editing: Marnie Holt, Dom McGrath, Tiarna McElligott; Project administration: Nantana Taptamat; Supervision: Dom McGrath, Marnie Holt.

References

Adarkwah, M. A., Amponsah, S., Huang, R., & Thomas, M. (2025). Artificial intelligence and human agency in education: Volume one: The nexus between AI and human agency in educational contexts. Springer.

Artino, A. R., & Jones, K. D. (2012). Exploring the complex relations between achievement emotions and self-regulated learning behaviors in online learning. *The Internet and Higher Education*, *15*(3), 170–175. https://doi.org/10.1016/j.iheduc.2012.01.006

ASCILITE TELAS. (n.d.). *The TELAS framework: A validated standards framework for evaluating the quality of online learning*. Retrieved October 12, 2025, from https://telas.edu.au/framework/

Bovill, C. (2020). Co-creation in learning and teaching: The case for a whole-class approach in higher education. *Higher Education*, *79*(6), 1023–1037. https://doi.org/10.1007/s10734-019-00453-w

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

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

Bretag, T. (Ed.). (2016*). Handbook of academic integrity*. Springer.

Broadbent, J., Panadero, E., Lodge, J. M., & de Barba, P. (2020). Technologies to enhance self-regulated learning in online and computer-mediated learning environments. In M. J. Bishop, E. Boling, J. Elen, & V. Svihla (Eds.), *Handbook of research in educational communications and technology* (pp. 37–52). Springer. https://doi.org/10.1007/978-3-030-36119-8_3

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies and academic achievement in online higher education learning environments: A systematic review*. The Internet and Higher Education*, *27*, 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*(1), 43. https://doi.org/10.1186/s41239-023-00411-8

Crawford, J., Cowling, M., Ashton-Hay, S., Kelder, J.-A., Middleton, R., & Wilson, G. S. (2023). Artificial intelligence and authorship editor policy: ChatGPT, Bard, Bing AI and beyond. *Journal of University Teaching and Learning Practice*, *20*(5). https://doi.org/10.53761/1.20.5.01

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science*, *4*, 50. https://doi.org/10.1186/1748-5908-4-50

Eaton, S. E. (Ed.). (2023). *Handbook of academic integrity* (2nd ed.). Springer.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, *1*(1). https://doi.org/10.1162/99608f92.8cd550d1

FoS Student Led-Observation for Course Improvement (SLOCI). (2023). Semester 2, 2023 Faculty of Science generative AI project final report. *Faculty of Science*, University of Queensland.

Gonsalves, C. (2025). Addressing student non-compliance in AI use declarations: Implications for academic integrity and assessment in higher education. *Assessment & Evaluation in Higher Education*, *50*(4), 592–606. https://doi.org/10.1080/02602938.2024.2415654

Hadwin, A. F., Järvelä, S., & Miller, M. (2011). Self-regulated, co-regulated, and socially shared regulation of learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 65–84). Routledge.

Healey, M., Flint, A., & Harrington, K. (2014). *Engagement through partnership: Students as partners in learning and teaching in higher education*. The Higher Education Academy.

Henderson, M., Bearman, M., Chung, J., Fawns, T., Buckingham Shum, S., Matthews, K. E., & de Mello Heredia, J. (2025). Comparing Generative AI and teacher feedback: student perceptions of usefulness and trustworthiness. *Assessment & Evaluation in Higher Education*, 1–16. https://doi.org/10.1080/02602938.2025.2502582

Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist*, *48*(1), 25–39. https://doi.org/10.1080/00461520.2012.748006

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Lodge, J. M., de Barba, P., & Broadbent, J. (2024). Learning with generative artificial intelligence within a network of co-regulation. *Journal of University Teaching & Learning Practice*, *20*(7), Article 02. https://doi.org/10.53761/1.20.7.02

Matthews, K. E. (2017). Five propositions for genuine students as partners practice. *International Journal for Students as Partners*, *1*(2), 1–9. https://doi.org/10.15173/ijsap.v1i2.3315

Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press.

McDonald, N., Johri, A., Ali, A., & Collier, A. H. (2025). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *Computers in Human Behavior: Artificial Humans*, *3*, 100121. https://doi.org/10.1016/j.chbah.2025.100121

McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research* (2nd ed.). Routledge.

Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology*, *106*(1), 121–131. https://doi.org/10.1037/a0033546

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16*(1). https://doi.org/10.1177/1609406917733847

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, 422. https://doi.org/10.3389/fpsyg.2017.00422

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Rahimi, M. (2025). Advocating synergising a critical dialogic approach with competencies-trained dialoguing with GenAI tools to enhance students' critical thinking and communication competencies. *Journal of University Teaching and Learning Practice*, *22*(2), 1-22. https://doi.org/10.53761/vaa8h214

Rana, V., Verhoeven, B., & Sharma, M. (2025). Generative AI in design thinking pedagogy: Enhancing creativity, critical thinking, and ethical reasoning in higher education. *Journal of University Teaching and Learning Practice*, *22*(4). https://doi.org/10.53761/tjse2f36

Rudolph, J., Ismail, M. F., & Popenici, S. (2024). Higher education's generative artificial intelligence paradox: The meaning of chatbot mania. *Journal of University Teaching and Learning Practice*, *21*(6), 14–48. https://doi.org/10.53761/54fs5e77

Smith, D., Sokoya, D., Moore, S., Okonkwo, C., Boyd, C., Lacey, M., & Francis, N. (2025). Embedding generative AI as a digital capability into a year-long skills program. Journal of University Teaching and Learning Practice, *22*(4), 1-24. https://doi.org/10.53761/fh6q4v89

Taptamat, N., Holt, M., McGrath, D., McElligott, T., Miller, L., & Purwanto, H. (2025). *From anxiety to agency: A case study in co-designing an ethical AI guide for science students*. In R. Fitzgerald (Ed.), *Inquiry in action: Using AI to reimagine learning and teaching*. The University of Queensland. https://doi.org/10.14264/74c5f7e

Tillmanns, T., Salomão Filho, A., Rudra, S., Weber, P., Dawitz, J., Wiersma, E., Dudenaite, D., & Reynolds, S. (2025). Mapping tomorrow's teaching and learning spaces: A systematic review on GenAI in higher education. *Trends in Higher Education*, *4*(1), 2. https://doi.org/10.3390/higheredu4010002

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, *53*(4), 5–23. https://doi.org/10.1007/bf02504682

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2