Your University Name

Department / School (e.g., School of Education)

AI-Augmented Writing Support for University Students: Effects on Learning, Academic Integrity, and Equity

— Thesis Cover Page —

Author: Antonio Hill

Master's Thesis (e.g., M.Ed., M.A., M.S.)

Supervisor: [Dr. First Last]

Submitted on November 06, 2025

The thesis is submitted in partial fulfillment of the requirements for the degree listed above. The work presented is my own, and any assistance or sources are acknowledged in accordance with university regulations.

© 2025 Antonio Hill. All rights reserved.

Abstract

Generative AI (GenAI) writing tools now sit inside the everyday workflow of university students, promising rapid feedback, clearer organization, and time savings—yet raising urgent questions about learning, integrity, and equity. This thesis investigates whether a structured, transparent model of AI-augmented writing improves authentic writing performance and why such improvements occur. Grounded in self-regulated learning, feedback literacy, and cognitive load theory, the study employs an explanatory sequential mixed-methods design across writing-intensive courses. The intervention asks students to use GenAI for critique and planning, document disclosure and process evidence, and convert feedback into targeted revisions; a business-as-usual condition serves as the comparator. Quantitative analyses estimate adjusted treatment effects and test mediation via self-regulation, feedback-use behaviors, and cognitive load, with moderation by prior achievement and equity-relevant subgroups. Qualitative interviews and focus groups explore mechanisms, enactment fidelity, and boundary conditions for responsible use. Anticipated contributions include an integrated explanation of how AI support enhances writing when it reduces extraneous load while preserving germane processing; design principles for assessment that align pedagogy and integrity; and practical guidance to ensure that benefits are distributed fairly across diverse learners. The thesis aims to move debates beyond prohibition versus permissiveness and toward evidence-informed, equity-attentive implementation.

Chapter 1: Introduction

In less than three years, generative AI (GenAI) writing tools have shifted from an intriguing novelty to a normalized component of students' study routines. Their most salient affordances—rapid idea generation, iterative text refinement, and on-demand formative feedback—map closely to the hardest phases of academic writing: planning, synthesizing sources, structuring arguments, and revising for clarity. Universities now confront a dual imperative. First, they must recognize that well-designed GenAI use can enhance learning by catalyzing self-regulated behaviors such as goal setting, metacognitive monitoring, and responsive revision. Second, they must protect the epistemic integrity of assessment so that grades continue to represent students' own knowledge and skill. International guidance has urged higher education to adopt a human-centered response rather than blanket prohibition, emphasizing capacity building, responsible use, and equity of access (UNESCO, 2023).

The sector's scale of adoption suggests that "wait and see" is no longer tenable. Recent national survey evidence indicates that the proportion of students reporting any use of AI tools has risen dramatically year-over-year, with text generation and editing among the most common purposes (HEPI, 2025). For institutions, this creates immediate pressure to stress-test assessments for robustness to AI assistance and to equip staff and students with concrete enactment strategies rather than vague proscriptions. At the same time, integrity infrastructures remain unsettled. Detection technologies can surface signals of AI-like text but are not determinative evidence and are known to vary in accuracy across document length and linguistic profiles; consequently, they must not serve as the sole basis for adjudicating misconduct (Turnitin, 2023, 2025). The stakes are therefore pedagogical as well as ethical: universities must

design learning environments in which transparent, educative uses of GenAI strengthen learning without eroding trust in assessment. (HEPI, 2025; Turnitin, 2023, 2025).

The scholarly conversation has expanded rapidly since late 2022. Early syntheses described opportunities and risks, arguing that large language models could scaffold complex writing and reasoning while cautioning against overreliance, hallucinated claims, and the possibility that students might outsource rather than develop disciplinary thinking (Kasneci et al., 2023). As empirical evidence accumulates, a more nuanced pattern is emerging. Mapping studies and systematic reviews in 2024–2025 report that when GenAI is positioned as a *support* for idea development, revision, and feedback uptake—rather than a wholesale replacement for student work—researchers observe improvements in engagement, confidence, and, in some contexts, writing performance (Lee, 2024; Tillmanns et al., 2025; Qi et al., 2025). Importantly, these effects appear to be mediated by how learners *use* feedback and regulate their writing processes, which aligns with established theories of self-regulated learning. (Kasneci et al., 2023; Lee, 2024; Tillmanns et al., 2025; Qi et al., 2025).

Policy-oriented work has, in parallel, foregrounded equity. International analyses stress that benefits and risks are unevenly distributed: access to reliable devices, digital and prompt-design literacies, institutional clarity about permissible uses, and culturally responsive teaching shape who gains and who is exposed to harm (OECD, 2024). This literature underscores that universities should address learning, integrity, and equity together rather than treating them as competing priorities. Where detection is used at all, recent sector commentary and reportage emphasize the need for caution because false positives—especially on short texts or among multilingual writers—can produce harm if uncritically operationalized (Wired, 2024;

Jisc/National Centre for AI, 2025). These contextual factors have direct implications for assessment design and student support. (OECD, 2024; Wired, 2024; Jisc, 2025).

Notwithstanding the speed of publications and guidance, three actionable gaps persist in both practice and research. First, many institutional policies articulate values and prohibitions but provide limited, operational guidance for *how* to embed GenAI during planning, drafting, and revising so that students internalize strategies rather than outsource cognition. While international guidance promotes human-centered, capability-building approaches, it leaves course-level mechanics—what to model in class, what to require in process evidence, how to stage disclosure—largely to local design (UNESCO, 2023). Second, the literature often treats learning, integrity, and equity as separate strands. Few studies model the *interactions* among these domains by, for example, testing whether AI-supported feedback practices mediate writing gains while also examining subgroup differences by prior preparation, language background, or disability accommodations (OECD, 2024; Lee, 2024). Third, evaluation is still dominated by average treatment effects with less attention to mechanisms. Recent meta-analytical work points to behavioral engagement gains but notes the need to specify mediators such as feedback literacy and self-regulation and to test boundary conditions with appropriate moderators (Qi et al., 2025; Tillmanns et al., 2025). Addressing these gaps requires designs that connect day-to-day tool use to durable learning, embed transparent integrity practices, and examine distributional effects across student groups. (UNESCO, 2023; OECD, 2024; Qi et al., 2025; Tillmanns et al., 2025; Lee, 2024).

A further practical gap concerns the interpretation of detection outputs. Technical notes from major vendors explicitly caution that AI-writing indicators should not be used as the sole evidence for adjudication and acknowledge contexts where accuracy is lower, such as short

submissions; nevertheless, implementations sometimes rely on these indicators as de facto arbiters of authorship (Turnitin, 2023, 2025; Wired, 2024). This disconnect between guidance and practice reinforces the need for integrated assessment redesign (e.g., process portfolios, oral defenses, and authentic tasks) that reduces the decision-making burden on detectors while promoting meaningful learning. (Turnitin, 2023, 2025; Wired, 2024).

The thesis proposes and tests a learning-centered, integrity-aware, and equity-attentive model for AI-augmented academic writing in higher education. Its primary purpose is to move beyond generalized claims about GenAI being "good" or "bad" for learning and instead specify *how* structured, transparent AI use can support the development of writing expertise without compromising authorship or fairness. The first aim is to estimate learning effects of a scaffolded GenAI-supported drafting and feedback sequence relative to business-as-usual instruction. The second aim is to explain mechanisms by testing theoretically motivated mediators—self-regulated learning practices and feedback-use behaviors—and assessing whether cognitive load is reduced in ways consistent with scaffolding rather than shortcutting. The third aim is to examine integrity and equity by evaluating alignment with institutional policy (including disclosure practices and process evidence) and by testing moderation effects across equity-relevant subgroups such as first-generation status, language background, and disability accommodations.

The significance of this work is twofold. Theoretically, it links self-regulated learning, feedback literacy, and cognitive load perspectives to contemporary AI practice, offering a coherent explanatory framework for when and for whom GenAI enhances writing. Practically, it yields actionable design principles for assignments, feedback cycles, and policy language—including guidance on when and how to combine process artifacts with product evaluations—so

that institutions can maintain confidence in what grades mean while expanding access to high-quality formative support (UNESCO, 2023; OECD, 2024; Tillmanns et al., 2025). By integrating quantitative outcome modeling with qualitative accounts of student and instructor experiences, the project responds to calls for context-sensitive, evidence-informed strategies that do not trade off inclusion for integrity or vice versa (HEPI, 2025; Jisc, 2025). (UNESCO, 2023; OECD, 2024; Tillmanns et al., 2025; HEPI, 2025; Jisc, 2025).

For analytical clarity, key constructs are defined as follows. Generative AI writing tools are large-language-model systems used to generate, transform, or critique text; in this study they function as scaffolded assistants during planning, drafting, revising, and reflecting. Self-regulated learning (SRL) refers to cyclical processes of planning, monitoring, and evaluating one's work; in writing, SRL includes purposeful prompt design, iterative revision in response to feedback, and metacognitive reflection on changes. Feedback literacy denotes learners' dispositions and capabilities to seek, make sense of, and act on feedback, whether human or AI-mediated, with an emphasis on closing the loop from insight to revision. Cognitive load is the mental effort invested in a task; effective design reduces extraneous load while preserving germane processing so that learners can devote attention to argument quality and evidence integration. Academic integrity in the AI era means adhering to transparent, permitted uses aligned with institutional policy and the assessment's purpose; AI-writing indicators, where used, are treated as signals to be interpreted with professional judgment rather than as verdicts (Turnitin, 2023, 2025). Equity concerns fair access to beneficial uses—including availability, literacy support, and protections against bias—so that marginalized groups are not disproportionately excluded or penalized (OECD, 2024). (Turnitin, 2023, 2025; OECD, 2024).

Chapter 2 (Literature Review) synthesizes empirical and policy literatures on GenAI and academic writing, integrating evidence on learning outcomes, integrity governance, and equity. It highlights unresolved tensions in assessment design, the mixed performance of detection technologies, and the need for capability building among staff and students (Kasneci et al., 2023; Lee, 2024; Tillmanns et al., 2025; OECD, 2024). Chapter 3 (Theoretical Framework & Hypotheses) integrates SRL, feedback literacy, cognitive load, and technology-use perspectives into a path model linking AI-supported practices to writing performance through mediators and moderators grounded in current evidence (Lee, 2024; Qi et al., 2025). Chapter 4 (Methodology) details an explanatory sequential mixed-methods design with quasi-experimental comparison, validated rubrics, and interviews/focus groups, along with ethical safeguards consistent with international guidance (UNESCO, 2023). Chapter 5 (Results) presents descriptive statistics, adjusted comparisons, mediation and moderation analyses, integrity indicators, and qualitative themes, followed by a cross-strand integration. Chapter 6 (Discussion) interprets findings relative to theory, reconciles benefits and risks, and foregrounds equity implications with practical boundary conditions. Chapter 7 (Practical Implications) translates results into course and policy design guidance, including concrete enactment steps for disclosure, process evidence, and assessment redesign. Chapter 8 (Limitations & Future Research) addresses validity, generalizability, and priority next studies. Chapter 9 (Conclusion) synthesizes contributions and offers a pragmatic pathway for responsible adoption. (Kasneci et al., 2023; Lee, 2024; Tillmanns et al., 2025; Qi et al., 2025; UNESCO, 2023; OECD, 2024).

Chapter 2: Literature Review

Search strategy & inclusion logic

The review synthesizes empirical and conceptual literature on generative AI (GenAI) and academic writing in higher education, with a focus on learning effects, integrity governance, and equity/access. Searches were conducted across Scopus, Web of Science, ERIC, and Google Scholar for publications from late 2022 to November 2025 using combinations of terms such as "generative AI," "large language model," "academic writing," "feedback," "self-regulated learning," "feedback literacy," "cognitive load," "integrity," "detection," "equity," "language inclusivity," and "disability support." Inclusion criteria prioritized peer-reviewed studies and authoritative policy analyses that reported clear methodology or conceptual frameworks relevant to writing and assessment in higher education. Commentaries were included only when they offered widely cited conceptual syntheses or sector guidance. Given the pace of change in 2023–2025, the review also considered sector reports providing prevalence data and practice-oriented implications when those reports were methodologically transparent and widely referenced in the field (e.g., national student surveys). Exclusion criteria removed purely technical model papers without educational application, non-tertiary contexts unless directly generalizable to higher education writing, and preprints with insufficient methodological detail. The final corpus integrates systematic reviews, quasi-experimental and experimental classroom studies, large survey reports, instrument development papers, and policy analyses. (UNESCO, 2023; OECD, 2024; HEPI, 2025).

Genai & Writing: Impacts On Comprehension, Transfer, Metacognition, Feedback Use

The post-2022 literature shows rapid movement from early commentary to empirical inquiry regarding GenAI's influence on academic writing. Early syntheses argued that large language models (LLMs) could scaffold challenging phases of writing—planning, organization, and revision—while also introducing risks such as hallucinated claims and the potential

displacement of metacognition by hyper-fluent, low-effort text (Kasneci et al., 2023). As studies with stronger designs emerged in 2024–2025, the field began reporting measurable effects on learning when GenAI is framed as a support for knowledge construction and feedback uptake rather than as a replacement for student work (Lee, 2024). Evidence indicates that LLM-enabled feedback, when aligned to rubric criteria and coupled with explicit student reflection, can improve writing quality and strengthen motivation and engagement (Meyer et al., 2024; Kinder et al., 2025). These results converge on a principle common to tutoring research: learner initiative and guided use matter as much as the tool's capabilities. (Kasneci et al., 2023; Lee, 2024; Meyer et al., 2024; Kinder et al., 2025).

Studies specifically targeting comprehension and transfer suggest that GenAI can assist sense-making during literature review and argument development when prompts cue students to justify claims, seek counterexamples, and cross-check sources. In classroom experiments where students iteratively requested explanations, critique, and alternative outlines from an LLM, researchers have reported improved structure and coherence in subsequent drafts; however, they caution that transfer beyond the immediate assignment depends on reflective activities that make strategies explicit (Lee, 2024; Kinder et al., 2025). Survey research across multiple programs shows that students most often use GenAI to explain concepts, summarize sources, and propose outlines before drafting, with a smaller but notable fraction inserting generated text verbatim; these patterns are strongly shaped by institutional guidance and course design (HEPI, 2025). Collectively, the literature indicates that comprehension gains and near-transfer to writing organization are most reliable when students use GenAI to interrogate ideas and plan revisions rather than to produce final text. (Lee, 2024; Kinder et al., 2025; HEPI, 2025).

Metacognition and feedback behaviors are central in this stream. Randomized and quasi-experimental studies have found that adaptive AI feedback—tailored to rubric dimensions and elicited through structured prompts—can increase revision depth and the likelihood that students act on feedback, effects that are mediated by perceived usefulness and self-efficacy (Meyer et al., 2024; Kinder et al., 2025). Exploratory designs in writing-intensive courses show that students who prompt GenAI to critique argument logic, evidence sufficiency, and audience fit tend to produce revisions that align more closely with rubric descriptors than students who request mere grammar corrections (Meyer et al., 2024). This pattern suggests that GenAI's value for metacognition lies not in surface-level edits but in the facilitation of evaluative judgment—what feedback literacy scholars would call making sense of quality and deciding how to act on it. When instructors require short "feedback-to-action" memos documenting how AI and human feedback were used and why, metacognitive benefits appear stronger and more stable across tasks (Lee, 2024; Hawkins et al., 2025). (Meyer et al., 2024; Kinder et al., 2025; Lee, 2024; Hawkins et al., 2025).

The literature also traces limits and risks. Several reviews in 2024–2025 caution that while GenAI can enhance engagement, the evidence for durable gains in complex disciplinary writing is still mixed due to short study horizons, small samples, and heterogeneous tasks (Feigerlova et al., 2025). Research in health and professional education similarly notes that improvements in short-term writing outcomes do not automatically translate to domain transfer, particularly when assignments demand specialized knowledge and citation accuracy (Zhui et al., 2024; Feigerlova et al., 2025). Some studies report that students over-trust confident but incorrect suggestions, highlighting the importance of explicit verification routines and triangulation with credible sources during literature synthesis. This is consistent with meta-

commentary arguing that educators should design prompts and checkpoints that require students to justify claims against source material rather than accept AI output at face value (Lee, 2024; Meyer et al., 2024). Overall, the weight of evidence suggests that GenAI can support comprehension, near-transfer, and metacognition when integrated as an interactive feedback partner and when students are required to document how feedback informed revision. (Feigerlova et al., 2025; Zhui et al., 2024; Lee, 2024; Meyer et al., 2024).

At the systems level, prevalence data shape how one interprets classroom effects. National-scale surveys show that student use has accelerated quickly, with the majority reporting GenAI use for assessment support and a substantive minority acknowledging direct insertion of AI-generated text (HEPI, 2025). These patterns underscore that studies of "if used" conditions in controlled settings may underestimate the real-world variability of use quality and disclosure practices. They also emphasize the need for instructional designs that channel widespread adoption toward productive behaviors and require process evidence to make students' learning visible (HEPI, 2025). Hepi+1

Self-Regulated Learning & Feedback Literacy With AI

A unifying theme across learning studies is the centrality of self-regulated learning (SRL) and feedback literacy as mechanisms through which GenAI affects writing. SRL frameworks emphasize goal setting, strategic help seeking, monitoring progress, and evaluating outcomes; when students use GenAI within this cycle to plan drafts, check logic, and evaluate revisions, they are more likely to realize gains in quality and confidence (Zimmerman, 2002; Lee, 2024). Recent studies that explicitly modeled SRL find that AI literacy and SRL together predict writing performance, with SRL exerting the stronger effect, suggesting that tool knowledge is insufficient without self-management of the writing process (Shi et al., 2025). Importantly, this

work shows that SRL is not a static trait: structured prompting protocols, revision journals, and guided reflection can cultivate SRL behaviors during AI-assisted writing. In contexts where instructors model prompt design and require transparent documentation of how feedback was applied, increases in self-monitoring and evaluative judgment are observed alongside improved rubric scores (Meyer et al., 2024; Hawkins et al., 2025). (Shi et al., 2025; Meyer et al., 2024; Hawkins et al., 2025; Lee, 2024).

Feedback literacy research further clarifies how students convert AI suggestions into learning. Studies in assessment and learning analytics show that learners who can interpret criteria, compare exemplars, and plan concrete revisions are better positioned to benefit from GenAI feedback than those who treat output as an answer key (Jin et al., 2025). Classroom experiments demonstrate that adaptive LLM feedback can strengthen both cognitive and affective-motivational outcomes, but only when paired with explicit activities that require students to explain edits, reconcile conflicting advice, and perform source checking (Meyer et al., 2024; Kinder et al., 2025). In simulated assessment tasks, students who framed prompts as requests for critique ("identify gaps in evidence," "stress-test my counterargument") produced more meaningful revisions than peers who requested surface corrections; this differential aligns with feedback literacy's emphasis on sense-making and actionable planning (Hawkins et al., 2025). These findings imply that instructors should scaffold not only *what* to ask GenAI but also *how* to evaluate and operationalize responses through mini-memos, checklists, and peer-review routines. (Jin et al., 2025; Meyer et al., 2024; Kinder et al., 2025; Hawkins et al., 2025).

There is also early evidence that SRL-oriented AI literacy courses can develop durable habits. Exploratory interventions in "AI and Writing" courses report gains in students' ability to plan and troubleshoot multi-prompt workflows, manage verification steps, and articulate

boundaries of permitted use (Anders, 2025). In these settings, metacognitive routines—such as maintaining a prompt log or pairing AI feedback with targeted human peer feedback—appear to inoculate against overreliance while retaining the efficiency benefits that students value. Conversely, when students adopt a passive stance, benefits attenuate and risks of superficial revision increase, highlighting the importance of classroom norms and explicit modeling. Importantly, SRL and feedback literacy are also equity levers: structured routines can reduce gaps in how effectively different groups use GenAI by making productive behaviors teachable and visible. (Anders, 2025).

Finally, SRL/feedback-focused work helps explain why prevalence alone tells us little about learning. National surveys document that most students use AI and many insert AI-generated text, yet the *quality* of use depends on whether students are encouraged to plan, monitor, and evaluate rather than to copy-paste (HEPI, 2025). Studies that align GenAI activities with SRL cycles and feedback literacy consistently report stronger learning signals, suggesting that design—not mere access—determines whether GenAI functions as a scaffold or a shortcut. This distinction lays groundwork for the methodological choices in the current thesis, which emphasize mediation by SRL and feedback behaviors and explicit documentation of process evidence. (HEPI, 2025; Shi et al., 2025; Meyer et al., 2024).

Cognitive Load Theory & Scaffolding Via AI (Benefits/Overreliance)

Cognitive Load Theory (CLT) provides a complementary lens for evaluating GenAI's role as cognitive scaffold versus cognitive crutch. CLT distinguishes between intrinsic load (task complexity), extraneous load (inefficient presentation), and germane load (schema construction). Properly designed, GenAI can reduce extraneous load by offering worked examples, reorganizing outlines, or rephrasing confusing passages; these affordances can free attention for

germane processing, such as integrating sources or evaluating evidence. Experimental and classroom studies reporting improved writing quality with adaptive LLM feedback are consistent with this account: students use AI to clarify criteria, identify missing warrants, or test counterarguments, thereby allocating more mental effort to higher-order moves (Meyer et al., 2024; Kinder et al., 2025). (Meyer et al., 2024; Kinder et al., 2025).

Yet CLT also predicts risks when assistance obscures core problem-solving steps. Over-scaffolding can suppress germane load if students accept AI output without engaging in explanation or self-explanation. Recent work developing an AI-assisted writing cognitive-load scale for L2 contexts identifies distinct load dimensions in human-AI collaborative writing, highlighting conditions where cognitive effort shifts from meaning-making to interface management or prompt tinkering (Fan et al., 2025). Parallel analyses in educational technology question uncritical applications of CLT and suggest that AI-driven adaptivity can both optimize and distort load depending on how learners interact with suggestions (Gkintoni et al., 2025; Patac et al., 2025). Inconsistent effects across domains likely reflect variations in scaffold quality and verification routines; when design requires justification, source checking, and reflective memos, learners appear more likely to transform assistance into learning. This body of work implies that load-aware GenAI integration should pair assistance with metacognitive checkpoints to prevent passive acceptance and ensure that cognitive effort remains aligned with the learning objective. (Fan et al., 2025; Gkintoni et al., 2025; Patac et al., 2025).

For the present thesis, CLT motivates measurement choices (e.g., perceived workload scales) and design principles (e.g., stepwise prompts that require students to verbalize rationales). It also supports a testable distinction between scaffolding and shortcutting: reductions in extraneous load should coincide with stable or increased evidence of germane processing (e.g.,

richer warranting and source integration), whereas shortcutting would show lower effort but brittle understanding. The mediation models proposed later leverage this distinction to test whether SRL and feedback use convert load savings into deeper learning.

Academic Integrity: Plagiarism/Contract Cheating, Detection Limits, Assessment Redesign

The integrity literature has evolved from initial alarm to cautious pragmatism. Sector surveys and journalism document high rates of GenAI use alongside inconsistent enforcement and uneven policy clarity at the course level. National-scale reports in the UK, for example, note a year-over-year increase from 66% to 92% in student use, with a minority directly inserting AI-generated text and many institutions encouraging staff to stress-test assessments rather than rely on detection (HEPI, 2025). Investigative reporting and institutional updates suggest that while misconduct cases exist, many universities have shifted toward policies emphasizing disclosure, process evidence, and assessment redesign to keep learning visible (The Guardian, 2025). These sources, while not peer-reviewed, are influential in shaping practice and echo concerns raised in policy analyses that simple prohibition is neither feasible nor educationally desirable. (HEPI, 2025; The Guardian, 2025).

Detection technologies represent a focal point of debate. Vendor technical notes and sector briefings emphasize that AI-writing indicators are probabilistic signals rather than verdicts; false positives can occur, particularly for short submissions or text with specific linguistic profiles (Turnitin, 2023). While some sector commentary in 2025 claims that false positive rates for mainstream detectors are “relatively low,” these communications acknowledge small-sample limitations and caution against sole reliance on detector outputs (Jisc/National Centre for AI, 2025). Journalistic coverage of detection at scale further complicates the picture, reporting large numbers of flagged papers alongside institutional hesitance to penalize without

corroborating evidence—an implicit recognition that detectors cannot shoulder the full burden of authorship judgments (Wired, 2024). The emerging consensus is that detectors may initiate conversations but should be embedded within broader integrity infrastructures that include transparent policy, disclosure norms, and assessment designs less vulnerable to substitution. (Turnitin, 2023; Jisc, 2025; Wired, 2024).

Assessment redesign is therefore a key integrity strategy. Policy guidance encourages educators to require process artifacts (e.g., annotated outlines, prompt logs, and revision histories), incorporate oral defenses or micro-vivas, and emphasize authentic, situated tasks where personal insight and course-specific data matter (UNESCO, 2023). Studies reporting stronger learning gains often pair AI feedback with these design elements, suggesting a synergy between integrity and pedagogy: the same features that make substitution more difficult also promote metacognition and transfer (Meyer et al., 2024; Kinder et al., 2025). Institutional adoption varies, but common trends include explicit statements of permitted use, disclosure requirements, and tiered consequences that distinguish misunderstanding from deception. From a research standpoint, there remains a need for longitudinal studies that examine how redesign affects both misconduct rates and learning outcomes across disciplinary contexts, and for evaluations that compare student experiences under different disclosure and process-evidence regimes. (UNESCO, 2023; Meyer et al., 2024; Kinder et al., 2025).

A related concern is the “silent normalization” of uncredited AI support. Survey data indicate that many students perceive AI as akin to spellcheck or grammar tools and do not recognize when disclosure is required, especially for paraphrase or idea generation (HEPI, 2025). Clearer policy language and classroom exemplars can reduce ambiguity by distinguishing conceptual support (e.g., brainstorming) from textual borrowing (e.g., inserting generated

passages) and by modeling how to cite or disclose appropriately. This aligns with feedback-literacy approaches that frame disclosure as part of reflective practice: students document what the tool contributed, why changes were made, and how they verified content. Designing writing assignments that reward transparent process and articulate consequences for misrepresentation appears to be more educative, and arguably more enforceable, than binary bans that are easy to circumvent. (HEPI, 2025; UNESCO, 2023).

Equity: Differential Access, Digital Literacy, Language Inclusivity, Disability Support

Equity analyses emphasize that GenAI's benefits and risks are unevenly distributed across student populations. Policy syntheses argue that access to reliable devices, high-quality internet, and paid tool features interacts with digital literacy and language background to shape outcomes, such that students with fewer resources may be less able to leverage advanced prompting, verification, or iterative revision (OECD, 2024). Without targeted supports, GenAI could widen existing achievement gaps by amplifying advantages for students who already possess strong academic literacies. Conversely, carefully scaffolded use can reduce barriers for multilingual writers and students with disabilities by offering immediate explanations, alternative phrasings, and accessible feedback formats. The equity case for GenAI is therefore not automatic; it is contingent on availability, literacy, and policy clarity. (OECD, 2024).

Within this literature, two lines stand out. The first concerns language inclusivity. Studies in L2 and multilingual writing suggest that AI-mediated paraphrase and explanation can lower entry barriers to disciplinary discourse by helping students test wording and examine register; however, researchers caution that uncritical use may mask misunderstanding or create overreliance on surface fluency (Fan et al., 2025). This risk is heightened when detectors produce higher false-positive rates for certain linguistic profiles, potentially chilling legitimate support

use if policies are not carefully framed (Wired, 2024). The second line concerns disability support. Policy papers and working reports argue that AI has potential to augment assistive technologies—for instance, by converting feedback into multimodal formats or by providing structure templates that reduce executive-function load—yet they stress the need for rigorous evaluation and co-design with disabled students to avoid paternalistic or one-size-fits-all solutions (OECD, 2025). Across both lines, equity turns on intentional design: access subsidies, AI-literacy training, and clearly signposted disclosure practices can transform GenAI from an inequity amplifier into an inclusion tool. (Fan et al., 2025; OECD, 2025; Wired, 2024).

A further equity consideration is the distribution of institutional training. National surveys report that while GenAI use is widespread, only a minority of students have received formal instruction on effective and ethical use. This asymmetry effectively creates an expertise gap: students who learn prompt design, verification, and disclosure in one course can outperform peers elsewhere, not because they are more capable writers but because they were taught the rules of engagement (HEPI, 2025). Policy guidance therefore recommends capacity-building for both staff and students, emphasizing that equitable benefit requires explicit teaching, not mere access (UNESCO, 2023; HEPI, 2025). In research terms, this implies that equity analyses should include both structural access and instructional exposure as predictors of outcomes and should test for interactions with subgroup characteristics such as first-generation status, language background, and disability accommodations. (HEPI, 2025; UNESCO, 2023).

Synthesis & Unresolved Tensions

Across literatures, three points of synthesis emerge. First, GenAI's educational value for academic writing is best understood as a function of designed interactions: when tools are positioned as feedback partners embedded in SRL cycles and paired with reflection and

verification, studies report improvements in writing quality, engagement, and confidence (Lee, 2024; Meyer et al., 2024; Kinder et al., 2025). Second, integrity and pedagogy are intertwined rather than opposed. The same assessment redesign features that make substitution harder—process evidence, oral defense, authentic tasks—also promote metacognition and transfer (UNESCO, 2023; Meyer et al., 2024). Third, equity is neither automatic nor incidental; it depends on access, literacy, and policy clarity. Structured training and disclosure norms can transform GenAI from a potential amplifier of existing disparities into a lever for inclusion, particularly for multilingual writers and students with disabilities (OECD, 2024; Fan et al., 2025). (Lee, 2024; Meyer et al., 2024; Kinder et al., 2025; UNESCO, 2023; OECD, 2024; Fan et al., 2025).

Nonetheless, important tensions remain unresolved. Evidence for durable transfer is still limited, and methodological heterogeneity makes meta-analytic aggregation challenging (Feigerlova et al., 2025). Detection technologies continue to raise fairness questions, especially in short-text contexts or among multilingual writers (Wired, 2024), and sector messaging about "low false positives" sits uneasily alongside cautions about small-sample validation (Jisc, 2025). Finally, equity interventions require resourcing and governance; without investment in training and access, widespread adoption could entrench disparities even as average outcomes rise. These tensions motivate the present thesis's mixed-methods design, which explicitly models SRL and feedback behaviors as mediators, tests subgroup moderation, and pairs quantitative outcomes with qualitative accounts of how students and instructors experience AI-supported writing in context. (Feigerlova et al., 2025; Wired, 2024; Jisc, 2025).

(The later part of the thesis has been intentionally omitted to protect data privacy)

Partial References List

Al-Zahrani, A. M. (2024). Exploring the impact of artificial intelligence on higher education in Saudi Arabia. *Humanities and Social Sciences Communications, 11*(1), Article 130. https://doi.org/10.1057/s41599-024-03432-4

Bauer, E., Wiesner, M., & Opitz, A. (2025). Effects of AI-generated adaptive feedback on statistical skills and interest. *British Journal of Educational Technology, 56*(5), e13609. https://doi.org/10.1111/bjet.13609

Francis, N. J., O'Neill, M., & Khan, R. (2025). Generative AI in higher education: Balancing innovation and integrity. *British Journal of Biomedical Science*, 82, 14048. https://doi.org/10.3389/bjbs.2024.14048

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

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). North-Holland.

Hawkins, B., Taylor-Griffiths, D., & Lodge, J. M. (2025). Exploring the effect of feedback literacy on AI-enhanced essay writing. *Assessment & Evaluation in Higher Education, 50*(4), 1–13. https://doi.org/10.1080/02602938.2025.2492070

Jisc National Centre for AI. (2023, September 18). *AI detection – Latest recommendations*. https://nationalcentreforai.jiscinvolve.org/wp/2023/09/18/ai-detection-latest-recommendations/

Jisc National Centre for AI. (2025, June 24). *AI detection and assessment – an update for 2025*. https://nationalcentreforai.jiscinvolve.org/wp/2025/06/24/ai-detection-assessment-2025/

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Kopp, B., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Computers & Education: Artificial Intelligence, 4*, 100188. https://doi.org/10.1016/j.caeai.2023.100188

Kinder, A., Wambsganss, T., & Kollar, I. (2025). Effects of adaptive feedback generated by a large language model: A case study in teacher education. *Computers & Education: Artificial Intelligence, 6*, 100352. https://doi.org/10.1016/j.caeai.2024.100352

Lee, D. (2024). The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. *Discover Education, 2*(1), 100225. https://doi.org/10.1016/j.desc.2024.100225

Lo, N., Chiu, Y.-H., & Lin, S. (2025). Evaluating teacher, AI, and hybrid feedback in university EFL writing. *SAGE Open, 15*(3), 21582440251352907. https://doi.org/10.1177/21582440251352907

Meyer, J., Seidel, T., & Rummel, N. (2024). Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases students' revision motivation and positive emotions. *Computers & Education: Artificial Intelligence, 5*, 100297. https://doi.org/10.1016/j.caeai.2024.100297

Organisation for Economic Co-operation and Development (OECD). (2024). *The potential impact of Artificial Intelligence on equity and inclusion in education* (OECD Artificial Intelligence Papers, No. 23). OECD Publishing. https://doi.org/10.1787/15df715b-en

Slimi, Z. (2023). The impact of artificial intelligence on higher education: Opportunities and challenges. *Education Journal, 12*(1), 1–10. https://files.eric.ed.gov/fulltext/EJ1384682.pdf

Sousa, A. E., Pinto, A., & Reis, L. P. (2025). Use of generative AI by higher education students: Adoption, practices, and concerns. *Electronics, 14*(7), 1258. https://doi.org/10.3390/electronics14071258

Sweller, J. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Turnitin. (2023, March 16). *Understanding false positives within our AI writing detection capabilities*. https://www.turnitin.com/blog/understanding-false-positives-within-our-ai-writing-detection-capabilities

Turnitin. (2025, October 14). *AI writing detection model: Release notes and guidance*. https://guides.turnitin.com/hc/en-us/articles/28294949544717-AI-writing-detection-model

Turnitin. (2025, July 15). *Guide for approaching AI-generated text*. https://www.turnitin.com/papers/guide-for-approaching-ai-generated-text

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

The Guardian. (2025, February 26). UK universities warned to “stress-test” assessments as 92% of students use AI. *The Guardian*. https://www.theguardian.com/education/2025/feb/26/uk-universities-warned-to-stress-test-assessments-as-92-of-students-use-ai

Wired. (2024, May 16). Students are likely writing millions of papers with AI, Turnitin says. *Wired*. https://www.wired.com/story/student-papers-generative-ai-turnitin

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