

Preserving Cognitive Ownership of Academic Writing in Higher Education: A Sustainable Hybrid Pedagogical Framework for Reasoning-Centred Artificial Intelligence Integration

¹Sunika Naz, ²Stella Sardar – ^{1,2}Arden University
³Imad Yasir Nawaz- Northumbria University, London

Abstract

This study explores how distinctive types of generative artificial intelligence (AI) practice, augmentation, co-construction, and replacement shape students' reasoning skills and sense of cognitive ownership in higher education (HE) academic writing. This research also responds to growing humanitarian concerns about the erosion of student commitment, the undermining of autonomy, and ethical learning in HE. To address this core gap, an explanatory sequential mixed-methods design was employed. Data were collected from 412 UK HE students, complemented with in-depth interviews from 24 participants. Quantitative modelling showed that augmentation strengthens reasoning through reflective engagement, co-construction yields mixed cognitive outcomes, and replacement significantly weakens ownership and efficacy. Qualitative findings revealed subsistent experiences behind these practices: some students articulated no ethical harm by AI-supported reflection, while others exhibited a quiet disarticulation of their self-learning skills. Incorporating these insights, this study proposed the Hybrid Human–AI Reasoning Integrity Model (HHARIM), a sustainable pedagogical framework in HE that centres human reasoning in ethical AI use. The recommended model also highlights cognitive ownership as an essential element and outlines a robust framework for responsible AI use to safeguard learning, ethics, and autonomy in HE. This study contributes theoretically by offering HHARIM as a framework for effectively embedding AI, thereby upholding ethical, sustainable, and human-centred learning. Ultimately, the implications of this proposed model will influence HE systems to encourage sustainable AI pedagogical practices that reinforce academic writing rather than compromise students' learning efficacy.

Keywords: Artificial Intelligence; Cognitive ownership; Sustainable Pedagogical; Higher Education.

Wordcount: 234

1.0 Introduction

The rapid evolution of generative Artificial Intelligence (AI) has initiated profound transformations in higher education (HE), fundamentally impacting traditional methods of student research, brainstorming, and academic writing. Since the emergence of large language models (LLMs) such as ChatGPT, Gemini, Claude, Grammarly and Copilot, these AI systems have become increasingly embedded in students' everyday learning practices, reshaping core academic behaviours related to critical thinking, assignment writing, and intellectual agency (Kasneji et al., 2023; Luckin, 2024; Zawacki-Richter et al., 2019). While universities have rigorous restrictions on plagiarism, authorship, and academic misconduct (Bretag, 2019; Dawson, 2020; Eaton, 2023), far less attention has been given to how AI influences cognitive ownership; the extent to which students retain responsibility and authorship over reasoning processes that assignment writing is designed to develop (UNESCO, 2023).

Academic writing has long served as a fundamental pedagogical tool for assessing knowledge in relevant areas, cultivating critical reasoning, and disciplinary identity (Bean, 2011; Wingate, 2012). Through processes of argument construction, evidence evaluation, explanation, and reflection, students develop the cognitive foundations essential to higher-order thinking (Andrews, 2010). However, the increasing availability of generative AI now challenges these assessments. AI's capacity to produce structured arguments, coherent explanations, and discipline-specific interpretations raises critical questions regarding what remains "human" in the reasoning demonstrated through written assignments (Selwyn, 2023; Williamson & Piattoeva, 2023). These concerns extend beyond textual originality to encompass deeper issues of human reasoning integrity, cognitive displacement, and intellectual authorship.

Emerging research suggests that AI can both enhance and abolish students' reasoning. Again, AI may provide scaffolding that strengthens metacognition, prompts reflection, and supports HE learners who struggle with structure or language (Holmes et al., 2022; Kasneci et al., 2023). However, over-reliance on AI-generated academic writing can lead to cognitive offloading, in which essential learning tasks are generated without human involvement, reducing opportunities for students to develop academic writing and critical thinking skills (Fischer et al., 2020; Yang, 2022). Existing literature and institutional policy, however, remain ambiguous and provide limited empirical clarity on how different modes of AI should be used, such as augmentation, co-construction, or replacement, to shape students' learning processes.

This significant gap in evidence is increasingly problematic as generative AI becomes normalised within higher education. Students often lack sufficient knowledge of the ethical boundaries between appropriate and inappropriate AI assistance (Cotton et al., 2023; Searle et al., 2024), and institutional guidelines rarely address the cognitive implications of AI-supported writing. UNESCO (2023) has called for governance models that prioritise human productivity and ethical decision-making, yet practical, pedagogically grounded frameworks remain scarce. There is an urgent need to reconceptualise academic integrity around reasoning integrity, the preservation of human-led cognitive processes within AI-supported writing, rather than focusing solely on textual originality.

To address these gaps, this study investigates: To what extent do different modes of AI use influence students' reasoning skills and cognitive authorship in academic writing, and how can HE integrate AI ethically while preserving human intellectual agency?

Drawing on a mixed-methods design, the study combines structural equation modelling (SEM) with qualitative thematic analysis to examine relationships between AI use, human reasoning performance, and authorship perceptions. In doing so, it foregrounds the cognitive dimensions of academic integrity and introduces the Hybrid Human–AI Reasoning Integrity Model (HHARIM), a framework designed to support ethical, reasoning-centred AI integration in higher education. By addressing the cognitive, ethical, and pedagogical implications of AI-supported writing, this study responds to a growing need for evidence-based strategies that leverage AI's educational potential while safeguarding students' knowledge and skill development.

Generally, this research aims to contribute, both practically and theoretically, by proposing HHARIM, a sustainable, pedagogical and governance-oriented model that encourages ethical, sustainable, and human-centred AI adoption in HE.

2.0 Literature Review

Academic writing has long been acknowledged as a core approach through which higher education assesses students' critical reasoning, skills, academic knowledge and disciplinary identity (Lee et al., 2024). Over the past two decades, research has emphasised that writing is not merely a communicative output but a complex cognitive process that enables interpretation, argumentation, and reflective judgement (Andrews, 2010; Lea & Street, 1998; Wingate, 2012). Through the iterative structure of drafting, reviewing and revising, students internalise the reasoning skills that underpin advanced academic competence (Bean, 2011). Yet, as generative AI tools accelerate rapidly, this pedagogical function is undergoing significant disruption.

Generative AI and Cognitive Transformation in HE

Generative AI systems particularly large language models (LLMs) are now widely used across HE for brainstorming, summarising texts, revising drafts, and transcribing conceptual explanations (Kasneci et al., 2023; Zawacki-Richter et al., 2019). Scholars have recognised that AI can provide linguistic scaffolding, personalised feedback, and cognitive support, potentially enhancing learning, especially for students facing linguistic or structural writing challenges (Holmes et al., 2022; Dwivedi et al., 2024). Moreover, research grounded in digital learning frameworks suggests that AI-enabled tools may improve learning efficacy, reduce writing anxiety, and broaden access to academic support (Thompson & Lee, 2012; Weller, 2020).

However, the transformative capacity of AI raises pedagogical and epistemic concerns. AI-generated text can resemble high-quality academic writing without requiring the underlying reasoning effort traditionally associated with scholarly work (Selwyn, 2023; Williamson & Piattoeva, 2023). This creates risks of cognitive offloading, where essential tasks such as conceptualisation, evidence evaluation and argument construction are transferred from the student to the AI system (Fischer et al., 2020; Sparrow & Chatman, 2019). As a

result, academic integrity concerns have expanded beyond issues of plagiarism and authorship to include the erosion of critical thinking skills and intellectual autonomy (Bretag, 2019; Eaton, 2023).

Academic Integrity Beyond Plagiarism: The Rise of Cognitive Integrity

Between 2000 and 2020, academic integrity discourse was dominated by concerns over cheating, plagiarism detection, and text-based misconduct (Dawson, 2020; Macfarlane et al., 2014). While these concerns remain relevant, scholars now argue that generative AI necessitates a broader conceptualisation of integrity, one that includes cognitive effort, reasoning transparency, and the human ownership of intellectual processes (Eaton, 2023; Harper et al., 2021). Moreover, UNESCO's (2023) global guidance echoes this shift by emphasising human agency as a non-negotiable foundation for AI use in HE.

This emerging discourse suggests that traditional plagiarism frameworks are insufficient for addressing AI's cognitive implications. When AI produces arguments, explanations, or conceptual connections on behalf of students, the issue is not simply one of textual authorship but of diminished cognitive ownership, a student's sense of control and responsibility over the reasoning embedded in their coursework (Luckin, 2024). Loss of cognitive ownership undermines the very purpose of academic writing as a process for developing critical thinking (Lea & Street, 1998; Wingate, 2012).

AI-Augmented, Co-Constructed, and AI-Generated Reasoning

Recent literature distinguishes between three primary modes of student–AI interaction: augmentation, co-construction, and replacement. Augmentation involves using AI as a scaffold to support human reasoning—for example, brainstorming ideas, generating outlines, clarifying concepts, or prompting metacognitive reflection. Existing Research shows that augmentation can enhance reasoning by stimulating cognitive engagement and enabling deeper reflection (Kasneci et al., 2023; Holmes et al., 2022).

Co-construction reflects iterative collaboration between human reasoning and AI-generated suggestions. While this mode can encourage active decision-making and comparative evaluation, scholars warn that constant back-and-forth can dilute cognitive effort, as learners may over-rely on AI prompts (Yang, 2022; Popenici & Kerr, 2017).

Replacement describes scenarios in which AI performs core reasoning tasks, such as interpreting data, generating arguments, or drafting paragraphs on behalf of students (Crompton & Burke, 2023). Studies repeatedly show that replacement undermines learning by reducing cognitive effort, weakening reasoning skills, and limiting the development of intellectual agency (Fischer et al., 2020; Selwyn, 2023; Williamson & Eynon, 2020).

Despite growing conceptual clarity, empirical research examining the relationship between AI use and human reasoning outcomes remains limited. Existing studies focus largely on user perceptions, ethical concerns, or plagiarism risks, rather than outlining how different patterns of AI engagement influence human reasoning performance or ownership (Cotton et al., 2023; Searle et al., 2024). This gap highlights the need for robust empirical evidence.

Cognitive Ownership as a Mediating Construct

While cognitive ownership is increasingly recognised in theoretical discussions, few studies have empirically measured it or examined its structural role in student learning (Bailey, 2023; Harper et al., 2021). The literature emphasises that high levels of ownership are associated with deeper cognitive engagement, greater metacognitive awareness, and stronger intellectual autonomy (Bandura, 2001; Zimmerman, 2002). Conversely, cognitive displacement through automation can weaken agency and long-term reasoning capability (Fischer et al., 2020). Recent theorists argue that AI's educational impact cannot be understood solely through performance outcomes; instead, it must be analysed through the lens of ownership and authorship of cognition (Luckin, 2024; Selwyn, 2023). This perspective positions cognitive ownership as a potential mechanism explaining why augmentation supports reasoning and why replacement undermines it.

However, the mediating role of cognitive ownership in AI-assisted writing remains underexplored empirically. The literature lacks mixed-methods studies that integrate quantitative measurement of reasoning with qualitative insights into students' lived cognitive experiences.

3.0 Ethical and Pedagogical Frameworks for AI Integration

The expanding presence of AI in education has prompted calls for governance frameworks that preserve human agency and support ethical decision-making (UNESCO, 2023; Eaton, 2023). Existing literature proposes that ethical AI use in HE should prioritise transparency, human-led reasoning, and pedagogical structures that discourage cognitive offloading (Holmes & Tuomi, 2022; Williamson & Eynon, 2020). Yet existing frameworks tend to focus on institutional policy rather than on the pedagogical or cognitive mechanisms that preserve the integrity of reasoning.

As Weller (2020) and Popenici and Kerr (2017) highlight, meaningful AI integration must include pedagogical safeguards such as reflective tasks, justification prompts, and metacognitive checkpoints to ensure students remain the primary originators of their thinking. Without these safeguards, AI risks transforming academic writing from a reasoning-centred practice into a productivity exercise detached from cognitive development. Additionally, UNESCO (2023) highlights the importance of placing human reasoning as the principle of creation. While AI can be an effective tool for proofreading, there are still only a few practical pedagogical models available.

After critically scrutinising the existing literature, the core emerging gap is identified: the absence of sustainable, pedagogically grounded models that pilot ethical, human-reasoning-centred AI integration in HE. This study addresses these gaps by empirically evaluating AI-use modes, modelling the role of cognitive ownership in reasoning performance, and proposing the Hybrid Human–AI Reasoning Integrity Model (HHARIM) as a framework for sustainable, ethical and human-centred AI adoption in higher education.

4.0 Conceptual Model Summary

The framework ascertains AI-use modes as the primary independent variable, human reasoning skills as the dependent variable, and cognitive ownership as the mediating mechanism that influences the relationship. Reasoning integrity frames the broader ethical context.

This conceptual grounding informed the following study hypotheses:

- H01: AI-use mode predicts reasoning skills.
- H02: Cognitive ownership mediates the relationship between AI-use mode and reasoning skills.
- H03: AI-use mode influences perceived academic integrity through a sequential pathway:

AI-use mode → cognitive ownership → human reasoning skills → integrity perceptions.

The conceptual logic underlying this study begins with the recognition that generative AI has become deeply embedded in academic writing, fundamentally altering how students approach reasoning, ownership, and knowledge-based work. As students interact with AI in different ways, two distinct modes of use emerge: augmentation and replacement, each carrying different cognitive implications. Augmentation supports students' thinking by prompting reflection, offering scaffolding, or generating examples, thereby reinforcing their cognitive ownership. In contrast, replacement allows AI to take over the reasoning process, producing arguments and explanations on the student's behalf, which risks diminishing the very reasoning skills academic writing is designed to develop. This presents a challenge to both reasoning integrity, the preservation of human-led cognitive processes, and cognitive ownership, the sense that students retain authorship over their knowledge construction. Because these risks arise not from the technology alone but from how it is used, effective responses require a combination of governance (clear rules, transparency, institutional oversight) and pedagogy (teaching strategies that foreground reasoning and metacognition). Together, these elements define safe, ethical, and academically sound boundaries for AI use. The Hybrid Human–AI Reasoning Integrity Model (HHARIM) embodies this logic by positioning human reasoning as the anchor of all AI-supported academic writing practices, ensuring that AI acts as a tool for knowledge gathering rather than a substitute for cognitive effort. In doing so, the framework promotes sustainable, ethical, and human-centred reasoning in academic practices in HE.

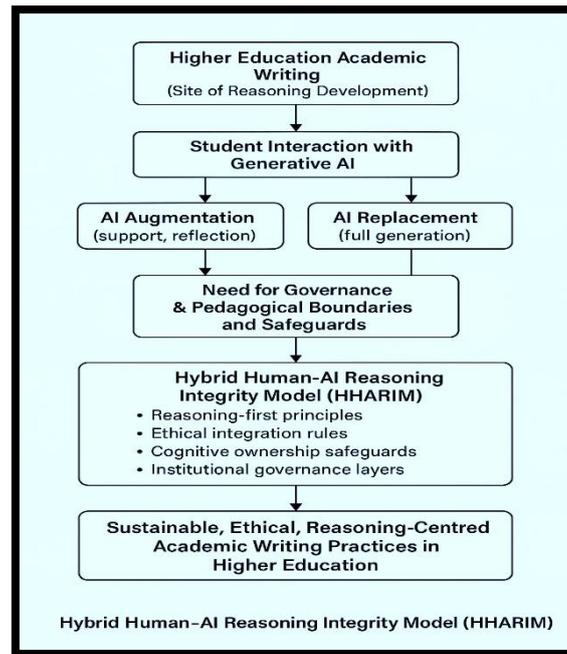


Figure 1: Proposed Conceptual Framework.

5.0. Methodology

This study adopted an explanatory sequential mixed-methods design, grounded in a pragmatic paradigm, to investigate the extent to which and how different generative AI-use modes influence students' reasoning and cognitive ownership. The quantitative phase involved 412 UK HE students recruited through voluntary sampling, using a survey measuring AI-use mode, reasoning skills (CRI), and a validated five-item cognitive ownership scale. Data were analysed using SPSS and AMOS SEM, enabling examination of group differences and mediating pathways. The qualitative phase purposively selected 24 diverse participants for semi-structured interviews exploring human-centred reasoning processes, ownership, and ethical boundaries in AI use. Thematic analysis, supported by NVivo, combined inductive and deductive coding, with credibility ensured through triangulation, peer debriefing, and an audit trail. Integration occurred through side-by-side comparison of themes and SEM pathways, informing the development of the Hybrid Human-AI Reasoning Integrity Model (HHARIM). Ethical approval and standard protocols were observed throughout.

The sample represented a broad mix of disciplines, academic years, and levels of AI familiarity. Participants provided information on their frequency and mode of AI use in academic writing, self-rated their reasoning ability, and completed the validated Critical Reasoning Inventory (CRI) (1–5 scale), which measures analytical thinking, argument construction, evidence evaluation, and reflective judgement.

- **AI-Use Mode:** Students selected statements corresponding to one of three theoretically derived categories—Augmentation, Co-construction, and Replacement—reflecting the degree to which AI supports, collaborates in, or substitutes for their reasoning.
- **Reasoning Skills (CRI Score):** Established measure assessing analytical and evaluative reasoning competencies.
- **Perceived Cognitive Ownership:** A five-item scale ($\alpha = .82$) developed for this study to assess the extent to which students believed their submitted writing represented their own thinking.
- **Control Variables:** Discipline (STEM vs. non-STEM), year of study, and prior academic writing experience.

Again, semi-structured interviews using maximum variation sampling based on AI-use intensity (high/low), discipline, and gender ensured that diverse user profiles were represented, particularly in terms of their reasoning approaches and experiences with AI-supported writing. Interviews explored:

- how students engaged with AI during their writing process,
- how they understood the boundary between their own reasoning and AI contributions,
- how they interpreted ownership of ideas and academic integrity.

Building on the literature that distinguishes among augmentation, co-construction, and replacement modes of AI use (Holmes et al., 2022; Selwyn, 2023) and their divergent effects on reasoning, and grounded in the HRF principle, which highlights human-led cognitive processes, this study tested three hypotheses. These hypotheses reflect mediation and sequential-mediation relationships that align with this study’s variables and conceptual framework.

6.0. Findings

Consistent with the explanatory sequential design, the quantitative results established the structural relationships among key variables, and the qualitative analysis subsequently enriched and explained these patterns. The integration of these two strands offers a nuanced account of students’ reasoning practices in the context of AI-supported writing. It directly informs the development of the Hybrid Human–AI Reasoning Integrity Model (HHARIM).

The quantitative phase (n = 412) revealed apparent differences in reasoning performance and cognitive ownership across the three AI-use modes. Augmentation was the most common mode (42%), followed by Co-construction (31%) and Replacement (27%). Mean CRI scores varied significantly: Augmentation users scored highest (3.73), Co-construction users scored moderately (3.39), and Replacement users scored lowest (2.94). ANOVA results confirmed significant differences across groups, $F(2,409) = 24.6, p < 0.001$, with Tukey tests revealing the hierarchy: Augmentation > Co-construction > Replacement. Cognitive ownership scores followed the same pattern, with augmentation users reporting the highest ownership (M = 4.12) and replacement users reporting the lowest (M = 2.68). Structural equation modelling ($\chi^2/df = 1.97, CFI = 0.96, RMSEA = 0.045$) demonstrated that AI-use mode significantly predicted reasoning skills ($\beta = 0.43, p < 0.001$) and perceived cognitive ownership ($\beta = 0.57, p < 0.001$). Reasoning skills also predicted ownership ($\beta = 0.51, p < 0.001$), supporting a sequential-mediation pathway in line with the study’s hypotheses. Control variables were non-significant.

Figures 2 and 3 show the distribution of AI-use modes among students: Augmentation (42%), Co-construction (31%), Replacement (27%). Figure 3 demonstrates the mean CRI reasoning scores for each AI-use mode: Augmentation = 3.73, Co-construction = 3.39, Replacement = 2.94.

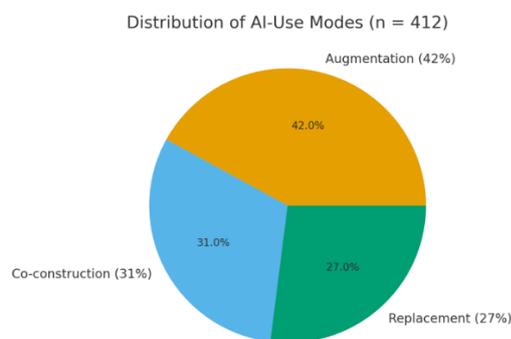


Figure 2: Distribution of AI-use modes (n=412).

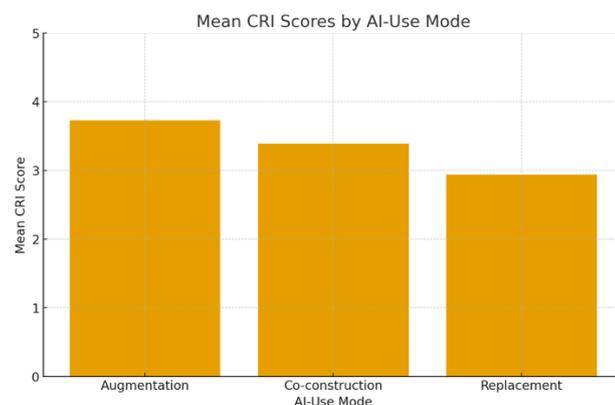


Figure 3: Mean CRI Scores by AI-use mode.

In addition to examining AI-use modes, the quantitative findings also revealed clear patterns in the specific AI tools students used for academic writing. As shown in Figure 4, students engage with a broad ecosystem of AI technologies, including both traditional text-enhancement tools and advanced generative AI systems. Grammarly (96 users) and ChatGPT (90 users) emerged as the most frequently used tools, followed by Microsoft Copilot (77 users), Google Gemini (75 users), and Quillbot (74 users). This highlights the pedagogical importance of understanding how different tools contribute to augmentation, co-construction, or replacement modes of AI use.

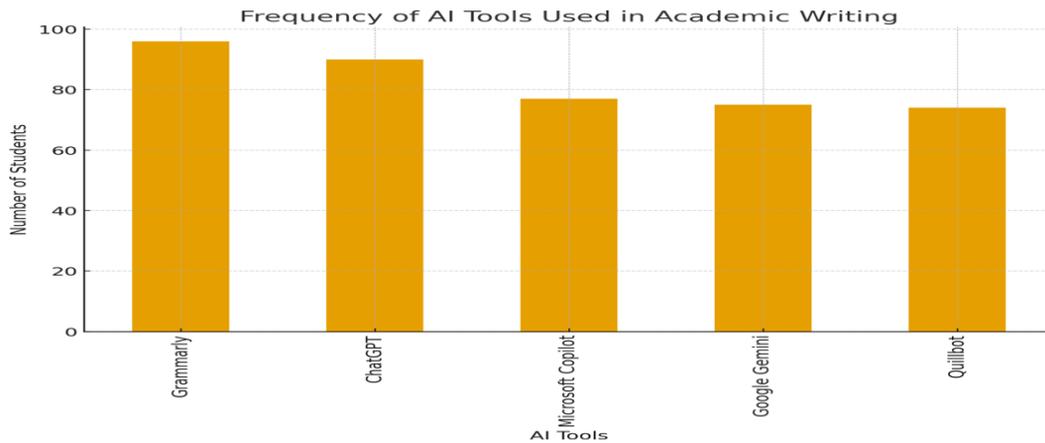


Figure 4: Frequency of AI tools used in academic writing

Reasoning Displacement (Explains Low Scores for Replacement Users)

Semi-structured interviews with 24 students elaborated and contextualised the statistical results. Four themes captured students' lived experiences of AI-supported writing and revealed nuanced mechanisms behind the quantitative relationships. Students who used AI in a replacement mode consistently reported that AI “took over” the cognitive work. Many described submitting text they did not fully understand: *“It sounded good, but if someone asked me to explain it, I wouldn’t know where to start.”* This theme helps explain the lower CRI scores and lower cognitive ownership observed among replacement users in the quantitative data. Participants explicitly linked reduced reasoning effort to reduced ownership, a pattern mirrored in the SEM pathway, which showed a negative association between replacement use and reasoning skills.

AI-Enabled Reflection (Explains Higher Scores for Augmentation Users)

Students who used AI augmentatively described AI as a reflective tool that supported idea generation, organisation, and metacognitive awareness: *“AI helps me notice gaps in my argument—it doesn’t replace my thinking but makes me think more clearly.”* This theme explains why Augmentation users showed the strongest CRI and ownership scores. It aligns with the SEM result, indicating a strong positive path from AI-use mode to reasoning skills when the use is supportive rather than substitutive.

Ethical Augmentation (Explains Strong Ownership Scores)

Many augmentation and co-construction users articulated an ethical commitment to preserving ownership of ideas. Students described internal boundaries that guided their use: *“If the ideas are mine and AI helps me express them better, I feel I’m still the author.”* This theme directly supports the mediation pathway (AI-use → ownership → reasoning) and reinforces the HRF principle: humans retain responsibility for reasoning, even when AI assists in expression.

Conditional Co-Construction (Explains Mid-Range CRI Scores)

Co-construction users described working interactively with AI while maintaining critical agency. These students challenged AI suggestions, compared alternatives, or integrated multiple versions: *“It’s like bouncing ideas around with something that responds, but I’m the one making the final choices.”* This theme explains why co-construction users scored moderately on CRI, higher than replacement but lower than augmentation. AI-supported reasoning, but the greater reliance on iterative cycles diluted some of the profound reasoning benefits seen in augmentation.

Table 1: Thematic Coding Framework.

Participant Statement	Code	Primary Theme	Main Theme	Proposition
“It sounds good, but I can’t justify it.”	Lack of reasoning ownership	Reasoning Displacement	AI Replacement Weakens Cognition	Replacement AI reduces reasoning depth
“AI helps me find gaps in my argument.”	Metacognitive prompting	AI-Enabled Reflection	Augmentation Strengthens Cognition	Augmentation improves reasoning
“If the ideas are mine, AI is just helping.”	Ethical self-regulation	Ethical Augmentation	Ownership as Integrity	Cognitive ownership predicts integrity
“AI gives options; I choose what to keep.”	Active human agency	Conditional Co-construction	Human–AI Agency Balance	Co-construction supports moderate reasoning
“I rely on AI too much sometimes.”	Cognitive offloading	Reasoning Displacement	AI Overdependence	Overreliance undermines learning
“AI helped me reorganise my paragraph.”	Structural support	AI-Enabled Reflection	Augmentation Benefits	Structural augmentation builds confidence

Integration: Explaining Convergence and Divergence

Table 2: Mixed Method Convergence Results.

Quantitative Finding	Qualitative Confirmation
Augmentation → highest reasoning	AI-enabled reflection strengthened cognitive processes.
Replacement → lowest reasoning	Reasoning displacement theme.
Augmentation → strongest ownership	Ethical augmentation practices.
AI-use → ownership → reasoning (mediation)	Students linked authorship with cognitive effort.

This convergence reinforces the validity of the HRF principle and underscores the centrality of ownership in reasoning-based academic integrity.

Table 3: Mixed Method Divergence Results.

Divergence	Quantitative Evidence	Qualitative Evidence	Interpretation & Implication
1. Overestimation of Ethical Confidence	Surveys showed high self-reported ethical confidence in AI use, suggesting students believed they were applying AI responsibly.	<ul style="list-style-type: none"> Interviews revealed uncertainty, inconsistent boundaries, and confusion about what constitutes ethical AI use. Students frequently said things like: <i>“I think it’s okay... but I’m not sure where the line is.”</i> 	<ul style="list-style-type: none"> Indicates that students overestimate their ethical competence. Highlighting a gap between perceived and actual ethical understanding. Strong need for clear institutional guidelines, training, and examples of acceptable AI use.
2. Perceived Learning vs. Actual Cognitive Engagement	Replacement users did not report significantly lower confidence or perceived learning in surveys.	<ul style="list-style-type: none"> Students admitted feeling more efficient but less cognitively engaged, e.g., <i>“I get work done faster, but I don’t think I’m actually learning.”</i> 	<ul style="list-style-type: none"> Shows a disconnect between perceived productivity and true reasoning development. Suggests risk of cognitive offloading not captured by quantitative measures. Institutions must address this with pedagogical interventions that promote active reasoning.

These variances highlight areas where institutional guidance and pedagogical intervention are required. Integrated findings show that AI’s effect on human reasoning depends not only on its integration but also on how it is engaged.

Augmentation highlights and reinforces human reasoning and sense of ownership.

Co-construction supports reasoning moderately.

Replacement undermines both reasoning and ownership.

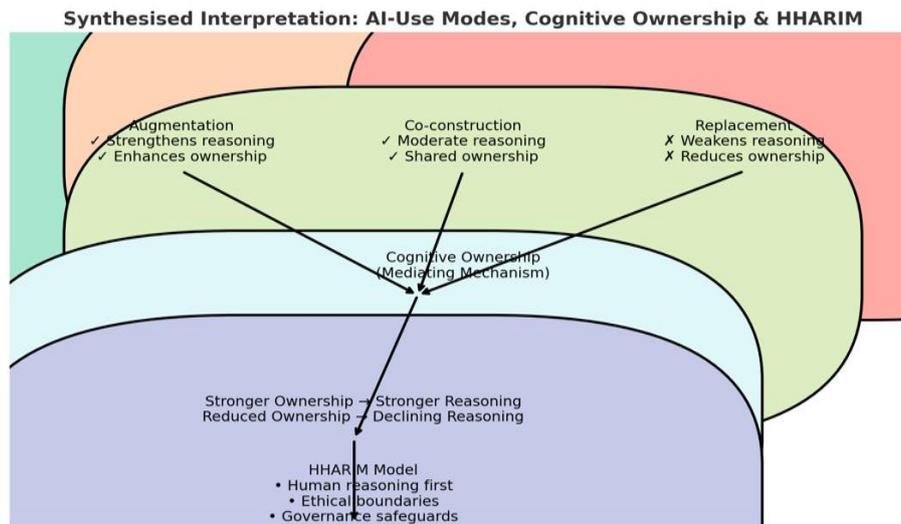


Figure 5: Synthesis Interpretation: AI-use Modes, Cognitive Ownership & HHARIM.

The SEM pathways and qualitative insights assemble on a core conclusion: cognitive ownership is the mechanism through which AI use influences human reasoning skills. In other words, when students feel that the work truly reflects their own skills and knowledge, their reasoning becomes stronger, more deliberate, and more ethically founded. When ownership diminishes, reasoning declines accordingly. These understandings directly formed the development of the Hybrid Human–AI Reasoning Integrity Model (HHARIM), which regards human reasoning as the essential foundation of academic writing and offers governance principles for HE.

7.0. Discussion

The findings directly support this inquiry by showing that the impact of AI on human reasoning depends not only on how it is used but also on the mode of engagement—Augmentation, Co-construction, or Replacement. This distinction clarifies an area where existing literature has mainly focused on plagiarism, textual misconduct, or general attitudes towards AI, while neglecting the cognitive effects of AI-supported writing processes (Selwyn, 2023; Cotton et al., 2023).

The literature review highlighted concerns about the displacement of cognitive effort when AI becomes a primary producer of content (Luckin, 2024; Williamson & Piattoeva, 2023), alongside emerging arguments that generative AI can enhance students’ reflective capacity when used critically (Kasneji et al., 2023). The present findings provide strong empirical support for both positions but show that the outcome depends on how students interact with AI. Augmentation users demonstrated the highest reasoning scores and most substantial ownership—aligning with work suggesting that AI can scaffold metacognition and enhance learning skills when used as a supportive tool rather than an automated generator (Holmes et al., 2022). Co-construction, while beneficial, yielded moderate support for human reasoning, consistent with the literature identifying productive human–AI collaboration but also warning of over-reliance (Yang, 2022). Conversely, replacement users disclosed weaker reasoning and diminished cognitive ownership, reinforcing concerns about cognitive offloading and loss of authorship (Fischer et al., 2020; MacFarlane, 2019).

Significantly, this investigation expands existing research by empirically demonstrating, through structural equation modelling, that cognitive ownership is the key mechanism connecting AI-use mode to reasoning outcomes. This offers a more precise cognitive explanation than previous work, which tended to frame AI risks mainly in ethical or integrity terms. The qualitative findings provide nuance, showing that even HE students who believe they are using AI ethically may unknowingly diminish their cognitive involvement when relying on AI-generated text. This divergence highlights the overestimation of ethical competence, a phenomenon that is not thoroughly addressed in the current literature.

The preliminary results of this paper strongly reinforce the Human Reasoning First (HRF) principle proposed in the literature review: ethical academic writing must centre on preserving human reasoning as the primary intellectual engine. When HE students use AI as a reflective tool or structural scaffold, their reasoning remains active, and they retain the key generators of ideas. However, when AI replaces cognitive

effort by generating arguments, explanations, or conceptual links, students lose control over their reasoning processes and experience lower levels of authorship. These findings sharpen the theoretical argument that reasoning integrity is the foundational component of academic integrity in the AI era. This position moves beyond traditional concerns with originality or plagiarism.

The relationship between human reasoning and ownership is not just correlational; it is sequential. The SEM findings and qualitative narratives confirm that greater ownership leads to stronger reasoning, while lesser ownership leads to weaker reasoning. This supports the idea that intellectual ownership is vital for deeper cognitive engagement (Bean, 2011). Therefore, the study confirms a reasoning-focused approach to academic integrity that emphasises human cognitive effort, not text production, as central to ethical AI use. The deviations identified in the mixed-methods analysis emphasise the challenges facing HE. However, students expressed high ethical confidence in surveys, yet interviews revealed uncertainty about the boundaries between ethical enhancement and unethical replacement. This gap suggests that students' internal ethical frameworks are inadequate for navigating the complexities of AI. As emphasised in the literature (Bretag, 2019; Eaton, 2023), academic integrity in modern contexts requires more than mere compliance; it needs explicit pedagogical guidance on human reasoning processes, authorship boundaries, and ethical judgment.

The preliminary findings suggest three key implications:

Institutions must provide more straightforward, reasoning-focused AI guidelines.

Generic warnings about plagiarism are insufficient. The students require explicit demonstrations of the difference between thinking with AI and thinking via AI.

Pedagogy should highlight cognitive ownership. Learning activities must require students to articulate their human reasoning steps, justify their decisions, and reflect on their thinking processes, practices that discourage the use of replacement-based AI.

AI literacy training must integrate ethical reasoning. Students need structured opportunities to critique AI outputs, identify illusions, and evaluate the epistemic reliability of generative AI models.

These implications are consistent with emerging calls in the literature for “critical AI literacy” and for academic institutions to develop governance frameworks tailored to student reasoning processes (Holmes & Tuomi, 2022).

8.0. Conclusion

This analysis explored how different modes of generative AI use—augmentation, co-construction, and replacement shape students' reasoning skills and cognitive ownership in higher education. Across both quantitative and qualitative approaches, a clear pattern emerged: AI's influence on learning is not determined by its existence but by how students engage with it. Augmentation enhanced human reasoning by supporting metacognitive reflection, co-construction offered moderate cognitive benefits, and replacement undermined both reasoning and perceived ownership. These findings provide strong empirical support for redefining academic integrity as human reasoning integrity, prioritising human-led cognition and intellectual autonomy over text originality alone.

A key theoretical contribution of this study is the identification of cognitive ownership as the central mechanism relating AI-use mode to reasoning outcomes. Structural equation modelling demonstrated a sequential pathway AI use influences ownership, ownership influences reasoning, and reasoning shapes perceptions of integrity. This clarifies why augmentation improves learning while replacement hinders it, filling a critical gap in the emerging literature on human–AI interaction.

The findings informed the development of the Hybrid Human–AI Reasoning Integrity Model (HHARIM), which offers a pedagogical and governance-focused framework for ethical AI integration. HHARIM positions human reasoning as the foundation of academic practices, sets boundaries for AI use, and emphasises the importance of human-reasoning-centred policies and AI literacy programmes. As most higher education institutions navigate the rapid normalisation of generative AI, this framework provides a practical and theoretically grounded guide to maintaining intellectual agency. However, this research also has limitations, including reliance on self-reported data, a cross-sectional design, and a sample drawn from UK institutions. Future research should employ longitudinal, experimental, and cross-cultural methods to examine how AI use patterns evolve over time and how they impact cognitive processes. Nonetheless, the findings highlight an important message: safeguarding human reasoning not merely detecting plagiarism is vital for fostering meaningful learning in the age of generative AI.

References

- Andrews, R. (2010) *Argumentation in higher education: Improving practice through theory and research*. Routledge.
- Bandura, A. (2001) 'Social cognitive theory: An agentic perspective', *Annual Review of Psychology*, 52(1), pp. 1–26. Available at: <https://doi.org/10.1146/annurev.psych.52.1.1>
- Bean, J.C. (2011) *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom*. 2nd edn. Jossey-Bass.
- Braun, V. and Clarke, V. (2021). 'One size fits all? What counts as quality practice in (reflexive) thematic analysis', *Qualitative Research in Psychology*, 18(3), pp. 328–352. Available at: <https://doi.org/10.1080/14780887.2020.1769238>
- Bretag, T. (2019) 'Academic integrity: Navigating the "new normal"', *International Journal for Educational Integrity*, 15(1), Article 1. Available at: <https://doi.org/10.1007/s40979-019-0040-9>
- Crompton, H. and Burke, D. (2023) 'Artificial intelligence in higher education: The state of the field', *International Journal of Educational Technology in Higher Education*, 20(1), Article 22. Available at: <https://doi.org/10.1186/s41239-023-00392-8>
- Cotton, D.R.E., Cotton, P.A. and Shipway, J.R. (2023) 'Chatting and cheating: Ensuring academic integrity in the era of ChatGPT', *Innovations in Education and Teaching International*, 61(2), pp. 228–239. Available at: <https://doi.org/10.1080/14703297.2023.2190148>
- Creswell, J.W. and Plano Clark, V.L. (2018) *Designing and conducting mixed methods research*. 3rd edn. SAGE.
- Dawson, P. (2020) 'Defending assessment security in a digital world: Preventing e-cheating and supporting academic integrity in higher education', *Higher Education Research & Development*, 39(1), pp. 44–58. Available at: <https://doi.org/10.1080/07294360.2019.1680950>
- Dwivedi, Y.K. et al. (2024) 'The future of generative AI: Academic perspectives on opportunities and challenges', *International Journal of Information Management*, 74, Article 102694. Available at: <https://doi.org/10.1016/j.ijinfomgt.2023.102694>
- Eaton, S.E. (2023) *Ethical AI and academic integrity in higher education*. University of Calgary Press.
- Fischer, K., Boone, G. and Neubert, J. (2020) 'Cognitive offloading and its implications for learning with digital technologies', *Educational Psychology Review*, 32, pp. 1–21. Available at: <https://doi.org/10.1007/s10648-019-09500-5>
- Harper, R., Bretag, T. and Rundle, K. (2021) 'Detecting contract cheating: Examining the role of assessment type', *Higher Education Research & Development*, 40(2), pp. 265–279. Available at: <https://doi.org/10.1080/07294360.2020.1724899>
- Holmes, W., Bialik, M. and Fadel, C. (2022) *Artificial intelligence in education: Promises and implications for teaching and learning*. 2nd edn. Centre for Curriculum Redesign.
- Holmes, W. and Tuomi, I. (2022) 'Critical AI literacy in higher education: A framework for teaching and learning', *AI & Society*, 37, pp. 1207–1221. Available at: <https://doi.org/10.1007/s00146-021-01270-5>
- Kasneci, E. et al. (2023) 'ChatGPT for education and research: Opportunities, threats, and strategies', *Frontiers in Education*, 8, Article 1197990. Available at: <https://doi.org/10.3389/educ.2023.1197990>
- Kline, R.B. (2016) *Principles and practice of structural equation modelling*. 4th edn. Guilford Press.
- Lea, M.R. and Street, B. (1998) 'Student writing in higher education: An academic literacies approach', *Studies in Higher Education*, 23(2), pp. 157–172. Available at: <https://doi.org/10.1080/03075079812331380364>
- Lee, D. et al. (2024) 'The impact of generative AI on higher education learning and teaching: A study of educators' perspectives', *Computers and Education: Artificial Intelligence*, 6, Article 100221. Available at: <https://doi.org/10.1016/j.caeai.2024.100221>
- Luckin, R. (2024) *AI for learning: Why artificial intelligence matters for education*. Routledge.
- Macfarlane, B., Zhang, J. and Pun, A. (2014) 'Academic integrity: A review of the literature', *Studies in Higher Education*, 39(2), pp. 339–358. Available at: <https://doi.org/10.1080/03075079.2012.709495>
- Morgan, D.L. (2007) 'Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods', *Journal of Mixed Methods Research*, 1(1), pp. 48–76. Available at: <https://doi.org/10.1177/2345678906292462>

- Popenici, S. and Kerr, S. (2017) 'Exploring the impact of artificial intelligence on teaching and learning in higher education', *Research and Practice in Technology Enhanced Learning*, 12(1), Article 22. Available at: <https://doi.org/10.1186/s41039-017-0062-8>
- Searle, A., Gašević, D. and Dawson, S. (2024) 'Understanding student use of generative AI: A learning analytics perspective', *Computers & Education*, 204, Article 104893. Available at: <https://doi.org/10.1016/j.compedu.2023.104893>
- Selwyn, N. (2023) *Should robots replace teachers? AI and the future of education*. 2nd edn. Polity Press.
- Shute, V.J. et al. (2021) 'The role of AI in supporting metacognition and learning', *Educational Psychologist*, 56(4), pp. 250–262. Available at: <https://doi.org/10.1080/00461520.2021.1956566>
- Sparrow, B., Liu, J. and Wegner, D.M. (2011) 'Google's effects on memory: Cognitive consequences of having information at our fingertips', *Science*, 333(6043), pp. 776–778. Available at: <https://doi.org/10.1126/science.1207745>
- Thompson, R. and Lee, M.J. (2012) 'Talking with students through screencasting: Experimentations with video feedback to improve student learning', *Journal of Interactive Technology and Pedagogy*, 1(1), pp. 1–16. Available at: <https://jitp.commons.gc.cuny.edu/talking-with-students-through-screencasting-experimentations-with-video-feedback-to-improve-student-learning/>
- UNESCO (2023) *Guidance for generative AI in education and research*. UNESCO Publishing.
- Weller, M. (2020) *25 years of ed tech*. Athabasca University Press.
- Williamson, B. and Eynon, R. (2020) 'Automating education: AI technologies and the governance of learning', *Learning, Media & Technology*, 45(1), pp. 1–6. Available at: <https://doi.org/10.1080/17439884.2019.1667823>
- Wingate, U. (2012) 'Using academic literacies and genre-based models for academic writing instruction: A "literacy" journey', *Journal of English for Academic Purposes*, 11(1), pp. 26–37. Available at: <https://doi.org/10.1016/j.jeap.2011.11.006>
- Yang, S. (2022) 'Human–AI collaboration in writing: Opportunities and cognitive risks', *Computers & Education*, 187, Article 104574. Available at: <https://doi.org/10.1016/j.compedu.2022.104574>
- Zimmerman, B.J. (2002) 'Becoming a self-regulated learner: An overview', *Theory into Practice*, 41(2), pp. 64–70. Available at: https://doi.org/10.1207/s15430421tip4102_2