

Human-AI Academic Co-Writing: A Systematic Review of Cognition, Authorship, and Responsibility

Vahid Goruslu

V.Garousi@qub.ac.uk

Queen's University Belfast

Systematic Review

Keywords: Academic writing, Artificial intelligence, Generative AI, Human–AI collaboration, Co-writing, Authorship, Cognitive offloading, Scholarly integrity

Posted Date: October 15th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11245863/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Generative artificial intelligence is increasingly involved in academic writing, yet existing reviews have largely examined isolated aspects such as writing outcomes, feedback, language learning, cognitive offloading, or editorial policy. This systematic literature review examines human–AI academic co-writing as an integrated phenomenon, with particular attention to the distribution of writing and intellectual labor, cognitive involvement, authorship, responsibility, and integrity. The review synthesizes 45 primary studies published between 2023 and 2026, including selected scholarly preprints that met the review's quality criteria, and situates them against 19 related secondary studies. Four questions guide the review: how the field has evolved across time and contexts; how human and AI roles are distributed across writing stages and with what benefits and trade-offs; how the literature conceptualizes cognitive and intellectual relationships between human writers and AI; and how co-writing affects authorship, intellectual contribution, ownership, responsibility, transparency, and scholarly integrity. The findings show a rapidly expanding and multidisciplinary literature in which AI participates across ideation, structuring, drafting, revision, and checking, while human judgement, verification, and accountability remain central. The review identifies a recurring tension between augmentation and substitution, as well as between efficiency gains and epistemic costs. It further argues that the key distinction is not simply between human-written and AI-written text, but between different configurations of human intellectual control. A literature-derived explanatory causal model is proposed to integrate these relationships and to support more contribution-sensitive approaches to disclosure, authorship, and governance.

1 INTRODUCTION

Generative artificial intelligence (GenAI), particularly large language models (LLMs), has rapidly become part of academic and scholarly writing. Researchers and students increasingly use AI systems for tasks ranging from language correction and stylistic revision to brainstorming, outlining, drafting, summarization, feedback, and reference checking. As these systems move beyond surface-level editing toward more substantive involvement in the development of arguments and texts, academic writing is becoming less clearly attributable to a single human writer and more appropriately understood, in many cases, as a form of human–AI co-writing.

This development matters because academic writing is not merely the production of grammatically acceptable text. Writing can involve idea formation, argument construction, evaluation of evidence, disciplinary judgement, revision, and reflection. When AI participates in these activities, questions arise not only about writing quality and efficiency, but also about cognition, intellectual contribution, authorship, ownership, responsibility, and scholarly integrity. The central issue therefore extends beyond whether AI improves writing. It concerns how intellectual and linguistic work is redistributed when humans and AI jointly contribute to the production of academic texts.

The existing literature reflects this breadth, but it is also fragmented. Some studies focus on effectiveness, writing quality, or productivity; others examine feedback, scaffolding, self-regulated

learning, or second-language writing. A different body of work addresses cognitive offloading, learner agency, authorship, disclosure, provenance, or editorial policy. Recent conceptual contributions further question whether authorship should depend on direct prose production, substantial intellectual contribution, final responsibility, or some combination of these. These strands are closely related, yet they are often examined separately.

The secondary literature has developed equally quickly. Nineteen relevant reviews published between 2023 and 2026 were identified for the present study. These reviews address topics including AI-assisted academic writing tools and effects, L2 and EFL writing, feedback, scaffolding, cognitive offloading, prompt engineering, assessment, editorial policies, authorship, and integrity. However, relatively little work has synthesized these issues around the broader phenomenon of human–AI academic co-writing itself—that is, around how writing and intellectual labour are distributed between humans and AI, and what this distribution means for cognition, authorship, responsibility, and governance.

A further reason for such synthesis is that human–AI co-writing occurs in substantially different contexts. In student writing, the process of producing a text may itself be part of the learning objective, so delegating planning, reasoning, or revision to AI may affect skill development and cognitive engagement. In researcher-oriented scholarly writing, by contrast, the primary goal is usually the reliable production and communication of knowledge, making questions of expertise, verification, provenance, and accountability more central. Treating these contexts as equivalent risks overlooking important differences in what counts as appropriate AI involvement.

This review therefore adopts human–AI academic co-writing as its central unit of analysis. It includes both empirical and conceptual work, as well as selected scholarly preprints that met the review’s quality criteria, because the field is developing rapidly and some of its most relevant contributions have not yet completed formal publication processes. Rather than examining AI use as a single variable, the review considers the stages at which AI participates, the kinds of cognitive and intellectual work delegated or retained, and the implications for authorship and responsibility.

The study is guided by four research questions:

- RQ1. How has research on human–AI academic co-writing evolved across time, disciplines, and writing contexts?
- RQ2. How are human and AI roles distributed across the academic writing process, and what benefits, limitations, and trade-offs are reported?
- RQ3. How does the literature conceptualize the cognitive and intellectual relationship between human writers and AI?
- RQ4. What implications does human–AI co-writing have for authorship, intellectual contribution, ownership, responsibility, transparency, and scholarly integrity?

By integrating these dimensions, the review aims to move beyond the question of whether AI is used in academic writing toward a more consequential question: how should academic co-writing be understood when cognitive, linguistic, and epistemic labor is distributed between human and artificial contributors?

2 RELATED REVIEWS

Research on generative AI in academic writing has already produced a substantial secondary literature. We identified 19 reviews published between 2023 and 2026 (see Appendix for the full list), indicating that this area has rapidly developed from an emerging topic into a comparatively well-reviewed field. However, these reviews tend to examine particular dimensions, educational settings, or disciplinary contexts rather than human–AI academic co-writing as an integrated phenomenon involving writing practices, cognition, intellectual contribution, authorship, and responsibility.

Several reviews take a relatively broad perspective on AI-assisted academic writing. Dergaa et al. (2023) provide an early synthesis of the prospects and potential threats of ChatGPT for academic writing, while Meng et al. (2026) systematically examine tools, measures, and reported effects. Dianelo et al. (2026) provide a cross-disciplinary scoping review of computational writing technologies in integrated learning environments. Sudirman et al. (2026) examine AI-assisted writing in educational institutions, and Neshkovska (2025) synthesizes reported merits and risks of AI-assisted academic writing. Collectively, these studies provide a broad understanding of how AI is being used in academic writing and of its reported benefits and limitations.

A substantial part of the review literature focuses more specifically on education and language learning. Jandildinov and Yersultanova (2024), Jandildinov et al. (2026), and Zadakhanova et al. (2025) examine AI-assisted scholarly writing in education, applied linguistics, and language education. Li (2026) focuses on higher-education EFL contexts, with particular attention to learning outcomes, self-regulated learning, and ethical governance. Roa and Halim (2024) and Tang et al. (2026) review AI-supported second-language writing, while Rahmat et al. (2026) examine AI as a scaffold for academic writing. Urzúa et al. (2025) concentrate specifically on AI-generated feedback and its effects on higher-education academic writing.

Other reviews address narrower dimensions of the emerging debate. Alghamdi and Alghizzi (2026) examine metacognitive filtering and cognitive offloading in AI-assisted L2 writing, bringing the cognitive consequences of AI use to the foreground. Xu et al. (2026) focus on the human role in AI-assisted writing assessment, Huang and Liu (2025) examine prompt engineering across AI-assisted writing domains, and Yoo (2025) reviews editorial policies defining acceptable boundaries of AI use in scientific writing. Asghar (2026) focuses on critical thinking and writing skills among university students. These reviews demonstrate that cognition, human agency, governance, and acceptable AI use have become important research concerns, although they are generally investigated separately.

The review most closely aligned with the present study is that of Slimi (2026), who critically synthesizes research on generative AI in relation to authorship, pedagogy, and academic integrity. This work brings together several concerns that are also central to human–AI academic co-writing. Nevertheless, the broader secondary literature remains distributed across relatively distinct strands concerned with effectiveness, pedagogy, language learning, feedback, assessment, cognition, prompt engineering, editorial policy, authorship, and integrity.

The present review therefore adopts human–AI academic co-writing as its central unit of analysis. Rather than treating AI primarily as a writing tool or examining one isolated consequence of its use, it considers how humans and AI jointly participate in producing academic and scholarly texts. This enables an integrated examination of four closely connected issues: the evolution and contexts of human–AI academic co-writing; the distribution of human and AI roles and associated benefits and trade-offs; the cognitive and intellectual relationship between human writers and AI; and the implications for authorship, intellectual contribution, ownership, responsibility, transparency, and scholarly integrity.

A further distinction is that the present review encompasses both student academic writing and researcher-oriented scholarly writing. Much of the existing secondary literature concentrates on higher education, EFL/L2 writing, learning, feedback, and assessment (Jandildinov & Yersultanova, 2024; Jandildinov et al., 2026; Zadakhanova et al., 2025; Urzúa et al., 2025; Li, 2026; Alghamdi & Alghizzi, 2026; Rahmat et al., 2026; Roa & Halim, 2024; Tang et al., 2026), whereas questions concerning scientific authorship, scholarly publication, and editorial governance are addressed in a smaller and partly separate body of work (Yoo, 2025; Slimi, 2026). Bringing these contexts together allows the review to examine whether assumptions about appropriate AI involvement differ when writing functions primarily as a means of learning and assessment versus as a means of producing and communicating scholarly knowledge.

Accordingly, the contribution of this review is not another general synthesis of the benefits and risks of AI-assisted academic writing. Instead, it integrates the emerging debate around co-writing itself: what humans and AI each contribute, how AI participation may affect human cognition and agency, and how these changes reshape established understandings of authorship and responsibility.

3 METHODOLOGY

This study employed a systematic literature review to synthesize research on human–AI academic co-writing. The review was designed to capture not only empirical evidence on the use and effects of generative AI in academic writing, but also conceptual, methodological, and normative contributions concerning cognition, authorship, intellectual contribution, responsibility, and scholarly integrity. The review therefore included both peer-reviewed publications and selected scholarly preprints that met the review’s quality criteria, provided that they met the relevance and quality criteria described below.

The search process combined database searching with backward and forward snowballing. Search terms were constructed around combinations of concepts such as generative AI, large language model,

ChatGPT, academic writing, scholarly writing, scientific writing, human–AI collaboration, co-writing, authorship, cognitive offloading, agency, and responsibility. In addition to database searches, the reference lists and cited-by records of relevant primary studies and 19 related secondary studies were examined to identify additional eligible publications. Duplicate records were removed before screening.

Studies were included when generative AI or large language models were substantively involved in the production, revision, feedback, evaluation, attribution, or governance of academic, scholarly, scientific, technical, thesis, or research-oriented writing, and when the study contributed to at least one of the review dimensions represented in the research questions. Eligible studies could be empirical, conceptual, theoretical, or methodological. Studies were excluded when they focused primarily on creative, professional, journalistic, or general-purpose writing without a substantial academic or scholarly writing component. Secondary studies were not included in the primary-study synthesis but were retained as related reviews and used for positioning and snowballing.

Because research on generative AI and academic writing is developing rapidly, publication status was not used as an automatic proxy for study quality. Peer-reviewed papers and scholarly preprints were both eligible. Preprints were included only when they presented a substantive scholarly contribution, had identifiable authors and institutional provenance, and provided sufficient methodological or conceptual transparency for quality appraisal. Where both a preprint and a later peer-reviewed version were available, only the published version was retained.

Quality assessment was conducted independently of publication status. Empirical studies were assessed with regard to clarity of aims, appropriateness of design, transparency of sampling and analysis, adequacy of evidence, and acknowledgement of limitations. Conceptual and theoretical studies were assessed in relation to clarity of argument, engagement with relevant literature, internal coherence, justification of claims, and contribution to the review questions. Methodological studies were assessed in terms of transparency, applicability, methodological justification, and limitations. Studies judged to provide insufficiently transparent or substantively weak evidence were excluded regardless of publication status.

For each included study, data were extracted on publication year, discipline and venue, study type, writing context, participant or document context where applicable, form of AI involvement, stages of the writing process addressed, and findings relevant to the four research questions. Particular attention was paid to the distribution of human and AI roles, reported benefits and trade-offs, forms of cognitive support or offloading, human agency and judgement, authorship and intellectual contribution, ownership, accountability, transparency, and scholarly integrity.

The final primary-study pool comprised 45 studies (P1–P45; see the Appendix for the full list), while 19 secondary reviews were retained as related work. The complete lists of primary and secondary studies are provided in the Appendix. The primary studies were synthesized using a combination of descriptive mapping and thematic analysis. Descriptive synthesis was used for publication trends, disciplinary distribution, writing contexts, and study characteristics. Thematic synthesis was then conducted in

direct alignment with the four research questions, with recurring concepts grouped into higher-level themes and areas of convergence, tension, and disagreement identified across the literature.

Detailed search strategies, study-selection information, eligibility criteria, quality-assessment procedures, data-extraction variables, and the complete study-level evidence and quality-appraisal matrix are provided in the *Supplementary Material*.

4 RESULTS

4.1 RQ1: Evolution and Characteristics of Human–AI Academic Co-Writing Research

The final primary-study pool comprised 45 studies published between 2023 and 2026. As shown in Fig. 1, publication activity increased steadily from 2023 to 2026, with the number of included studies rising from 5 in 2023 to 21 in 2026. The sharpest increase occurred in 2026, which accounts for nearly half of the 45-study corpus. This pattern indicates that human–AI academic co-writing has developed rapidly over a short period and remains a strongly expanding research area. The complete primary-study set is provided in the Appendix.

The disciplinary distribution was notably broad. The corpus included work from academic writing and language education, higher education, scholarly communication and publishing, research ethics and integrity, philosophy and ethics of technology, medical and health sciences, technical communication, computing and human–AI interaction, and specialist scientific and engineering domains. This diversity is visible, for example, in studies examining academic writing among university and language learners, authorship policies in scholarly journals, ethical guidelines for scientific writing, manuscript authorship and editorial processes in medicine, AI-supported technical communication, and AI-generated technical writing in electromagnetics. The phenomenon of human–AI co-writing is therefore not confined to a single disciplinary community; rather, different fields approach it through different concerns, ranging from learning and writing performance to authorship, provenance, accountability, and reliability.

The included studies also covered several distinct writing contexts. A substantial group focused on student academic writing, including undergraduate writing, second-language academic writing, teacher education, and the development of academic-writing skills. A second group examined thesis, dissertation, and researcher-oriented writing, including the experiences of researchers using AI to produce or revise scholarly texts. A third group addressed scientific and scholarly publication directly, including manuscript preparation, review-article writing, authorship policies, editorial governance, and questions concerning who should receive credit and assume responsibility when AI contributes substantially to a publication. Several studies crossed these boundaries, particularly those concerned with academic writing for publication or with frameworks intended to apply broadly across scholarly writing.

The corpus was also heterogeneous in study type and methodological orientation. Empirical studies included surveys, interviews, mixed-methods designs, experiments and quasi-experiments, text and interaction analyses, policy analyses, and Delphi-based consensus studies. Other publications were conceptual, theoretical, methodological, or normative contributions. These examined issues such as the meaning of authorship, substantial human contribution, distributed responsibility, provenance, ownership, and the attribution of human and AI contributions. Methodological studies further proposed concrete mechanisms for documenting or estimating the respective roles of human and AI contributors across stages of writing. Four included studies were scholarly preprints, retained because they directly addressed central questions of contribution and authorship and satisfied the review's quality criteria.

Taken together, RQ1 shows that the literature has expanded not only quantitatively but also conceptually and disciplinarily. The corpus suggests that earlier studies frequently centered on the immediate opportunities and risks of ChatGPT for academic and scientific writing, whereas the more recent literature increasingly addresses finer-grained questions concerning human–AI collaboration, cognitive involvement, intellectual ownership, provenance, responsibility, and the boundaries of legitimate authorship. This broadening of focus provides the basis for examining in the following sections how human and AI roles are distributed, how co-writing affects cognition and agency, and how established concepts of authorship and responsibility are being reconsidered.

4.2 RQ2: Human–AI Roles, Benefits, Limitations, and Trade-Offs

Across the primary studies, human–AI academic co-writing emerged less as a simple replacement of human writing by AI than as a division of labour across different stages of the writing process. AI was used for brainstorming, outlining, drafting, language refinement, restructuring, summarization, feedback, checking, and revision, while humans more consistently retained responsibility for selecting ideas, evaluating outputs, verifying claims and references, deciding whether revisions should be accepted, and taking responsibility for the final text. This staged distribution of work is made particularly explicit in the SLICE framework [P35], which distinguishes ideation, structuring, drafting, directive revision, implementation revision, checking, vetting, and guaranteeing. Similar patterns of hybrid participation are evident in studies of AI-supported academic authoring infrastructure [P1], scholarly publication workflows [P20, P30], scientific review writing [P5, P42], and student academic writing [P4, P6, P28, P39, P40].

The resulting division of labor is summarized in Table 2, which maps the main writing stages to the typical roles assumed by AI and human writers across the included studies. As Table 2 shows, AI involvement is distributed across nearly all writing stages, but the balance of responsibility shifts toward the human writer as the process moves from generation and revision toward verification, vetting, and final endorsement. This pattern supports the view that human–AI co-writing is best understood as a staged and asymmetric collaboration rather than as a uniform sharing of writing tasks.

Table 2
Human and AI roles across stages of academic co-writing

Writing stage	Typical AI role	Typical human role	Representative studies
Ideation	Generate possibilities, prompts, arguments	Select, reject, contextualize ideas	P2, P12, P25, P35
Structuring	Suggest outlines and organization	Determine argumentative structure	P1, P20, P35
Drafting	Generate prose or sections	Specify direction and assess substance	P12, P16, P30, P42
Revision	Rephrase, reorganize, improve style	Diagnose needed changes and approve revisions	P1, P6, P30, P39
Checking	Detect inconsistencies, language problems	Verify facts, references, reasoning	P5, P13, P16, P42
Vetting/guaranteeing	Limited role	Final judgement and accountability	P8, P26, P35, P38, P45

The most consistently reported benefit of AI assistance was increased efficiency, particularly in producing initial drafts, reformulating text, improving language, and accelerating revision. Studies involving scientific and technical writing show that AI can reduce the effort required to generate or reorganize prose [P1, P5, P13, P16, P42]. In student contexts, AI support was also associated with improvements in writing performance, organization, language accuracy, and perceived writing support [P4, P6, P39, P40]. For multilingual writers and researchers, AI could additionally reduce linguistic barriers and support clearer expression [P37]. These findings suggest that the value of AI is often greatest where the task is relatively well specified and where the human writer is able to evaluate and refine the generated output.

However, efficiency gains were frequently accompanied by additional verification and correction costs. AI-generated text could contain factual inaccuracies, invented or inappropriate references, oversimplified explanations, and stylistically convincing but substantively weak claims [P5, P13, P16, P21, P42]. In scientific and technical domains, these problems were especially consequential because fluent prose could mask technical errors. The comparison of human-only, AI-only, and AI-assisted scientific review writing in P42, for example, illustrates that AI assistance may accelerate text generation while simultaneously creating substantial fact-checking and editing work. Thus, faster drafting does not necessarily translate directly into lower total effort.

A second trade-off concerns quality versus homogenization and loss of authorial distinctiveness. AI-supported writing can improve surface quality, coherence, and readability, yet several studies raise concerns about originality, ownership, and the tendency toward standardized or generic expression [P7, P25, P37]. These concerns are closely connected to the degree of human involvement. Where AI is used

primarily for language support or implementation of clearly specified revisions, human control remains relatively strong. Where AI contributes extensively to ideation, argument construction, or full-draft generation, determining which intellectual elements originate with the human writer becomes more difficult [P2, P12, P25, P35, P45].

Human oversight therefore appears as a central condition of effective co-writing rather than an optional final check. Several studies emphasize active verification, critical evaluation, and explicit human responsibility for the final product [P8, P26, P38, P45]. The literature consequently points toward a hybrid model of co-writing in which AI may perform substantial generative and revisional work, but human contribution remains particularly important in judgement-intensive activities: defining purpose, selecting among alternatives, evaluating evidence, rejecting inappropriate outputs, integrating disciplinary knowledge, and guaranteeing the final manuscript. The principal trade-off is therefore not simply between “human writing” and “AI writing,” but between the amount of work delegated to AI and the amount of human effort required to supervise, verify, and intellectually endorse that delegation.

4.3 RQ3: Cognitive and Intellectual Relationships Between Human Writers and AI

The literature presents several distinct ways of conceptualizing the cognitive relationship between writers and generative AI. At the least intensive level, AI is treated as a tool or assistant that performs relatively mechanical or linguistic tasks such as editing, reformulation, grammar correction, and stylistic improvement [P1, P8, P13, P26]. In this configuration, the principal intellectual work remains clearly with the human writer, while AI reduces the effort required to express or refine already-developed ideas.

A second perspective presents AI as a cognitive scaffold or augmentative partner. Here, AI contributes not only to expression but also to planning, brainstorming, feedback, comparison of alternatives, and the iterative development of ideas [P18, P20, P26, P27, P29, P30]. Studies of students and researchers suggest that interaction with AI can prompt reflection, provide alternative formulations, stimulate idea development, and assist users in moving through difficult stages of writing. From this perspective, AI does not merely externalize existing thoughts; it can participate in the process through which those thoughts are clarified and developed.

Across the studies, these relationships can be organized into several recurring modes of cognitive involvement, summarized in Table 3. Table 3 highlights that the literature does not treat AI involvement as a single cognitive phenomenon. Rather, it ranges from relatively low-level assistance to forms of scaffolding, augmentation, and cognitive offloading, with increasing concern about reduced independent engagement as more substantive cognitive work is delegated to AI.

Table 3
Cognitive relationships in human–AI academic co-writing

Relationship	Characterization	Main opportunity	Main concern	Representative studies
Tool/assistant	AI supports expression and editing	Efficiency and language improvement	Limited intellectual concern	P1, P8, P13, P26
Scaffold	AI supports planning, feedback, reflection	Learning and idea development	Reliance on external support	P18, P20, P29, P36
Augmentation	Human and AI iteratively develop content	Expanded cognitive capability	Blurred contribution boundaries	P27, P30, P35
Cognitive offloading	Cognitive work is delegated to AI	Reduced effort/cognitive load	Reduced independent competence	P32, P39, P40
Cognitive substitution/dependence	AI performs substantial reasoning/writing work	High productivity	Loss of agency or intellectual engagement	P6, P12, P32

At the same time, the literature identifies the possibility of cognitive offloading, whereby activities previously performed internally by the writer are transferred to AI. P32 addresses this issue directly by distinguishing performance achieved with AI support from subsequent independent performance without AI. Other studies similarly raise concerns that extensive dependence on AI for planning, argument generation, drafting, or revision may reduce opportunities for writers to exercise the underlying cognitive skills themselves [P6, P29, P36, P39, P40]. Cognitive offloading is therefore not inherently negative: delegating lower-level operations may release attention for higher-level reasoning, but excessive delegation may also weaken engagement with precisely those processes through which academic-writing competence develops.

This tension makes human agency and judgement particularly important. Several studies portray productive AI use not as accepting generated material, but as continuously evaluating, selecting, rejecting, revising, and contextualizing it [P18, P26, P27, P29, P37]. The intellectual contribution of the human writer may therefore reside increasingly in directing the process and exercising judgement over alternatives rather than in personally producing every sentence. This interpretation is especially important for researcher-oriented scholarly writing, where the distinction between generating text and endorsing its intellectual content becomes central [P2, P12, P35, P45].

The literature consequently exposes a deeper disagreement about the relationship between writing and thinking. One position treats writing as partly executable: if the human provides the idea, argument, or conceptual direction, AI may generate substantial portions of the prose without necessarily displacing

human authorship [P12, P30]. An opposing concern is that writing itself is an important cognitive process through which arguments are discovered, refined, and tested; outsourcing substantial portions of writing may therefore also outsource part of the thinking [P32, P39]. The resulting question is not simply whether AI writes the words, but whether the human remains sufficiently engaged in the cognitive processes that give the words their intellectual meaning.

Overall, the findings suggest a continuum rather than a binary distinction between human and AI cognition. At one end, AI functions mainly as a linguistic tool; moving along the continuum, it becomes a scaffold, an augmentative collaborator, and potentially a substitute for portions of the writer's cognitive work. The consequences depend on what is delegated and on the degree to which the human continues to exercise metacognitive monitoring, disciplinary judgement, verification, and control. This relationship between cognitive participation and intellectual contribution also provides an important bridge to RQ4, because questions about authorship and responsibility ultimately depend partly on what cognitive and intellectual work the human writer continues to perform.

4.4 RQ4: Authorship, Intellectual Contribution, Responsibility, and Integrity

Across the corpus, authorship emerged as one of the most contested consequences of human–AI academic co-writing. A recurring point of agreement is that authorship cannot be reduced to the physical production of words. Several studies distinguish between text generation and the broader intellectual functions associated with authorship, including conception, argument development, critical revision, verification, approval, and responsibility [P2, P9, P11, P12, P26, P33, P41, P45]. This distinction becomes especially important when AI contributes not merely to language polishing but to ideation, structuring, argument generation, and drafting.

The literature nevertheless differs on what level of human contribution is sufficient for authorship. Some studies defend a relatively conservative position, maintaining that human authorship requires substantive intellectual contribution and cannot be preserved merely through final approval or responsibility [P11, P26]. Others argue that accepted academic practice already recognizes authors who may contribute little or no prose but who provide substantial intellectual direction, supervision, analysis, or critical judgement [P12]. The question therefore shifts from who wrote the sentences to who made the substantive intellectual contribution. P2 makes this issue explicit by proposing that substantial human contribution should be assessed through factors such as originality, skill, well-directed effort, and the strength of the connection between the human's intellectual input and the final manuscript. P35 similarly treats contribution as stage-specific, attempting to estimate human and AI participation across ideation, structuring, drafting, revision, checking, vetting, and guaranteeing.

A second major theme concerns the relationship between contribution and responsibility. Much of the literature maintains that AI cannot currently satisfy prevailing academic authorship criteria because it cannot meaningfully approve a manuscript, assume moral or institutional responsibility, or respond to challenges concerning the integrity of the work [P8, P26, P38]. Cheng et al. [P8], for example, emphasize

that scientific authorship combines contribution with accountability and that generative AI cannot guarantee the accuracy or integrity of a manuscript. P26 similarly places human vetting and guaranteeing at the center of responsible LLM use in academic writing. At the same time, newer conceptual work questions whether productive contribution and responsibility must remain bundled within one authorship role. P9 proposes distinguishing between contributing authors, who generate intellectual content, and certifying authors, who critically review, approve, and take responsibility for the work as a whole. This exposes a deeper tension in current authorship models: AI may participate substantially in generating intellectual content while being unable to bear responsibility, whereas a human may bear full responsibility without having produced much of the substantive content.

Questions of ownership and originality are similarly complicated by co-writing. Several studies suggest that originality cannot be assessed solely by textual novelty because the more important issue may be the provenance of ideas, arguments, and decisions [P7, P25, P31, P34]. AI-generated prose may be stylistically original while drawing opaquely on patterns learned from prior texts; conversely, a human may contribute the central ideas and argumentative structure while AI generates much of the final wording. This creates a provenance problem: readers may know that AI was used without knowing which parts of the intellectual work originated with the human and which were generated or transformed by AI [P34, P35]. The problem is therefore not simply plagiarism in the conventional sense, but the traceability of intellectual contribution.

This leads directly to the issue of disclosure. The literature broadly supports transparency but disagrees on how much disclosure is sufficient [P8, P19, P26, P34, P38, P45]. Minimal declarations such as “AI was used in preparing this manuscript” provide little information about the actual division of labor. More informative approaches describe the specific tasks performed by AI, the stages at which it was used, and the forms of human verification and control. P35 goes further by proposing a structured way of documenting stage-level contributions. The emerging direction is therefore from binary disclosure—AI used or not used—toward process-level transparency about who or what contributed to which parts of the scholarly work.

These issues converge around governance. Existing guidance tends to preserve three principles: substantial human intellectual contribution, human verification of AI-generated content, and human accountability for the final work [P8, P26, P38]. Yet the more extensively AI participates in reasoning and writing, the more difficult it becomes to apply conventional authorship criteria without modification. The literature therefore increasingly moves from asking whether AI should be called an author toward a broader question: how should academic credit, responsibility, and transparency be allocated when intellectual work is distributed across humans and AI systems?

The literature therefore does not converge on a single model of authorship and responsibility. Instead, several competing positions emerge concerning the basis of authorship, the relationship between contribution and accountability, the meaning of originality, and the level of transparency required. These positions are summarized in Table 4.

As Table 4 illustrates, the central disagreements concern not whether human responsibility remains necessary, but what type and degree of human intellectual involvement are sufficient to justify authorship. The literature is also moving beyond simple declarations of AI use toward more detailed accounts of provenance and task-level contribution. This suggests that future governance may need to distinguish more clearly between textual production, intellectual contribution, verification, and final accountability.

Table 4
Emerging positions on authorship and responsibility in human–AI academic co-writing

Issue	Position 1	Position 2	Representative studies
Basis of authorship	Authorship requires substantial human intellectual contribution	Authorship can persist without direct prose production if intellectual direction and judgement are substantial	P2, P11, P12, P26
Contribution vs responsibility	Contribution and accountability should remain linked	Productive contribution and certification/accountability may be separated	P9, P12
Originality	Originality concerns textual novelty and plagiarism	Originality also concerns provenance of ideas and intellectual decisions	P7, P25, P34
Disclosure	General disclosure of AI use may be sufficient	Stage- or task-specific disclosure is preferable	P8, P19, P26, P34, P35
Governance	Preserve existing human authorship criteria	Adapt authorship and contributorship models for distributed human–AI work	P9, P12, P33, P45

4.5 Integrative Synthesis: A Literature-derived Explanatory Model of Human–AI Academic Co-Writing

The findings across the four research questions suggest that the major dimensions of human–AI academic co-writing are not independent. Rather, the way writing tasks are distributed between humans and AI is connected to the degree of human cognitive involvement, the attribution of intellectual contribution, and the mechanisms required to maintain authorship, accountability, and scholarly integrity. Figure 2 integrates these relationships into a literature-derived explanatory causal model.

The model is conceptual rather than statistical. The positive and negative signs indicate the direction of relationships suggested by the reviewed literature and should not be interpreted as estimated causal coefficients, effect sizes, or quantitatively validated relationships. RQ1 provides the contextual layer of the model, including disciplinary setting, writing context, and type of scholarly activity, whereas the causal structure is derived primarily from the findings of RQ2–RQ4.

At the core of the model is the degree of AI task delegation. As more ideation, structuring, drafting, revision, or checking is delegated to AI, direct human cognitive involvement may decrease, particularly when AI use shifts from support or scaffolding toward offloading or substitution. Reduced direct involvement may, in turn, make the human intellectual contribution more difficult to identify and increase the importance of documenting contribution provenance.

At the same time, greater AI delegation increases the need for human judgement, verification, and transparent reporting. Strong human judgement and verification strengthen epistemic control over the resulting text and support more defensible claims to authorship and accountability. Similarly, clearer provenance and stronger governance mechanisms make it easier to determine who contributed what, who verified the work, and who can legitimately stand behind the final scholarly product.

The model therefore does not imply that greater AI involvement necessarily produces weaker authorship. Rather, it suggests that as AI assumes a larger share of writing and cognitive tasks, the conditions required for defensible human authorship become more demanding. Human judgement, verification, contribution traceability, and transparent reporting consequently become increasingly important.

5 DISCUSSION

The integrative model presented in Fig. 2 highlights the central finding of this review: the consequences of AI use depend less on whether AI is present than on how cognitive, intellectual, and epistemic labor is distributed between human and AI participants. Across the reviewed studies, human–AI academic co-writing emerges not as a single practice but as a range of configurations in which different writing and reasoning tasks are delegated, retained, supervised, or verified to different degrees.

One clear area of agreement is that human judgement remains central to defensible co-writing. Across educational, scientific, and conceptual studies, responsible AI use is consistently associated with human involvement in defining goals, evaluating outputs, verifying evidence, rejecting unsuitable suggestions, and taking responsibility for the final work [P8, P18, P26, P29, P38, P45]. This helps connect the findings across the review questions. The division of labor identified in RQ2 is not neutral: delegating more work to AI can change the degree of human cognitive engagement identified in RQ3, which in turn affects how intellectual contribution, authorship, and responsibility are understood under RQ4.

A central tension is therefore between augmentation and substitution. The same AI capability may support human thinking or replace part of it, depending on how it is used. Brainstorming with AI may stimulate reflection, but it may also substitute for idea generation; AI-supported revision may promote metacognitive evaluation, but it may also become passive acceptance; automated drafting may free time for higher-level reasoning, but it may also remove one of the processes through which reasoning itself develops [P18, P27, P29, P32, P39, P40]. The literature therefore does not support a simple conclusion that greater AI involvement is either beneficial or harmful. A more meaningful question is which cognitive functions are delegated and whether human judgement remains substantively engaged.

A related insight is that the relevant boundary is not between human-written and AI-written text, but between different configurations of human intellectual control. Two manuscripts may contain similar amounts of AI-generated prose yet differ substantially in the nature of human contribution. In one case, the human may provide the core argument, direct the drafting process, critically revise the text, verify the evidence, and stand behind the final manuscript. In another, the human may provide only a broad prompt and accept most of the resulting content. Treating both simply as “AI-assisted writing” obscures an important distinction. Studies such as P2 and P35 are particularly valuable because they shift attention from mere tool use toward the substance and traceability of contribution.

The review also reveals a tension between efficiency and epistemic cost. AI can accelerate drafting, editing, and language refinement, but these gains may be offset by the effort required for verification, correction, and reference checking [P5, P13, P16, P42]. In scholarly writing, faster text production has limited value if the resulting material requires substantial expert checking. This suggests that future evaluations of AI-assisted writing should consider total workflow cost, rather than measuring efficiency only in terms of generation speed.

An important distinction also emerges between student academic writing and researcher-oriented scholarly writing. In student writing, the process of producing the text is often itself part of the learning objective. Delegating planning, argument construction, or revision to AI may therefore weaken the development of skills that the assignment is intended to cultivate [P6, P29, P32, P36, P39, P40]. In scholarly publication, by contrast, the primary objective is generally the production and communication of reliable knowledge. Here, extensive AI assistance may be less problematic if the human researcher retains sufficient intellectual contribution, verification, disciplinary judgement, and responsibility for the work [P12, P26, P38, P42, P45]. The same degree of AI involvement may therefore be appropriate in one context and problematic in another.

This distinction has direct implications for governance. Policies based only on whether AI was used are too coarse. More useful governance would consider the purpose of the writing, stage of AI involvement, degree of intellectual delegation, level of human verification, and accountability structure surrounding the final product. The literature increasingly points toward stage-sensitive and contribution-sensitive governance rather than blanket permission or prohibition [P19, P26, P34, P35, P38].

The findings also suggest that current authorship models may be approaching a structural limit. Traditional academic authorship bundles together several functions: creating intellectual content, writing or revising text, approving the final version, receiving credit, and accepting responsibility. Human–AI co-writing makes it possible to separate these functions more sharply than before. AI may contribute to ideation and drafting without being able to assume responsibility, while humans may provide direction, verification, and certification without producing most of the prose. The proposals in P9 and P12 illustrate that future models may need to distinguish more explicitly between productive intellectual contribution and accountable certification. Whether such a separation should be institutionalized remains contested, but the underlying problem is likely to persist.

A further implication concerns transparency. The review suggests that binary disclosure is inadequate. A statement that AI was used does not reveal whether it corrected grammar, proposed an outline, generated an argument, drafted entire sections, or supplied references. If disclosure is intended to help readers assess intellectual contribution and responsibility, then it should provide information about the nature, stage, and extent of AI involvement [P26, P34, P35]. This does not necessarily imply publishing complete prompt histories, but it supports more informative contribution statements and clearer provenance reporting.

The review has several practical implications. For researchers, responsible co-writing requires retaining sufficient intellectual control to evaluate and defend the resulting work. For educators, the important question is not simply whether students use AI, but whether AI use supports or bypasses the cognitive processes that the task is intended to develop. For journals and institutions, disclosure policies should become more contribution-sensitive, and authorship guidance should address cases in which text production, intellectual contribution, and responsibility no longer coincide neatly. For tool designers, greater support for traceability, version histories, provenance, and transparent interaction logs could make responsible co-writing easier to document.

The review also has limitations. The literature is developing rapidly, and the study pool may become outdated relatively quickly. The corpus combines empirical, conceptual, methodological, and normative studies, which increases conceptual breadth but limits direct comparability. Several high-quality preprints were included because of the speed of development in this area; although they were subjected to quality assessment, their conclusions had not necessarily undergone formal peer review. The available literature is also uneven across disciplines and writing contexts, with stronger representation from higher education, language learning, and scholarly publishing than from some other domains.

Future research should move beyond self-reported perceptions of AI usefulness and increasingly examine actual co-writing processes. Interaction logs, version histories, prompt records, and longitudinal designs could reveal how human and AI contributions change across stages of writing. Particularly valuable would be studies linking process data with outcomes, such as whether extensive AI drafting affects later unaided writing ability, whether certain forms of delegation reduce independent reasoning, and whether stage-specific human oversight improves factual reliability. Stronger comparative work is also needed across student assignments, theses, and publishable research manuscripts.

6 CONCLUSION

This review examined human–AI academic co-writing as an emerging form of distributed scholarly work rather than merely as the use of AI writing tools. Across the 45 primary studies, the literature shows that AI can contribute across multiple stages of academic writing, from ideation and structuring to drafting, revision, and checking, while humans remain especially important in judgement-intensive activities such as verification, selection, endorsement, and accountability.

The findings also show that the central issues are deeply interconnected. The distribution of writing tasks affects the cognitive relationship between human and AI; this, in turn, shapes questions of intellectual contribution, authorship, ownership, provenance, and responsibility. The most important distinction is therefore not between human-written and AI-written text, but between different degrees and forms of human intellectual control.

Human–AI academic co-writing is likely to become increasingly common, but its legitimacy will depend on more than disclosure of AI use. It will require clearer understanding of who contributes what, how cognitive and intellectual work is distributed, and who is able to verify and stand behind the final scholarly product. Future research should therefore focus more directly on the processes of co-writing and on how different patterns of delegation affect learning, epistemic reliability, authorship, and accountability.

AI Use Disclosure Statement. OpenAI ChatGPT (GPT-5.6 Sol) was used to assist with literature organization, drafting, language refinement, and structural revision of the manuscript. The author critically reviewed, revised, and verified the resulting content and assumes full responsibility for the final manuscript.

Declarations

Author Contribution

This is a single-author paper. The single author has made all the contributions.

Acknowledgement

N/A

References

1. Abernethy, N. J. (2025). Let stochastic parrots squawk: Why academic journals should allow large language models to coauthor articles. *AI and Ethics*, 5(5), 4535–4553.
2. Alghamdi, L. H., & Alghizzi, T. M. (2026). Metacognitive Filtering and Cognitive Offloading in AI-Assisted L2 Writing: A PRISMA Guided Process-Tracing Synthesis. *Behavioral Sciences*, 16.
3. Alqahtani, Y. M., Alenazi, M. E., Alqahtani, M. A., Youcefi, N., & Irfan, M. (2026). AI-Assisted Writing and the Originality Problem: Evidence from Universities in the United Kingdom and Irish Universities. *journal of social sciences review*.
4. Alrefaie, Z., Alhazimi, A., Almarabheh, A. J., Madkhali, T., & Elsamanoudy, A. Z. (2026). AI assisted, mentor-guided narrative review writing task for medical students, a novel educational strategy to enhance research and academic writing. *Medical Teacher*, 48, 1012 – 1021.

5. Alyasin, A., & Shah, M. A. (2026). Human-AI Collaboration in Academic Writing: Exploring University Students' Agency through a Sociocultural Lens. *Ampersand*.
6. Amos, J. (2024). Ethical and effortful: workshopping human and generative AI academic writing collaborations. *Journal of Learning Development in Higher Education*.
7. Asghar, J. (2026). Constructing Legitimacy in AI-Assisted Academic Writing: Responsibility, Limitation, and Disclosure in Higher Education. *Global Social Sciences Review*, XI, 195–206.
8. Asghar, M. Z. (2026). Artificial Intelligence and Academic Writing: Exploring the Impact of AI-Assisted Tools on Critical Thinking and Writing Skills among University Students. *TAMSAAL*.
9. Awosanya, O. D., Harris, A., Creecy, A., Qiao, X., Toepp, A. J., McCune, T., Kacena, M. A., & Ozanne, M. V. (2024). The Utility of AI in Writing a Scientific Review Article on the Impacts of COVID-19 on Musculoskeletal Health. *Current Osteoporosis Reports*, 22, 146 – 151.
10. bin Johari, M. W., Khalili, N. N. M., Zainol, N. Z. B., & Nasaruddin, N. (2026). AI-Assisted Writing: Exploring ChatGPT's Influence on Undergraduate Writing Processes. *International Journal of Research and Innovation in Social Science*.
11. Butarbutar, R., & González Vallejo, R. (2025). Factors Influencing AI-Assisted Thesis Writing in University: A Pull-Push-Mooring Theory Narrative Inquiry Study. *Data and Metadata*.
12. Cao, L., & Lin, A. (2025). Dialogic triad: Balancing human-AI collaboration in academic writing for scholarly publication. *Journal of English for Research Publication Purposes*, 6(2), 340–369.
13. Chen, C., & Gong, Y. (2025). The Role of AI-Assisted Learning in Academic Writing: A Mixed-Methods Study on Chinese as a Second Language Students. *Education Sciences*.
14. Chen, X.-R. (2026). Layer-sensitive cognitive offloading in generative AI-assisted writing: supported performance and independent no-AI outcomes. *Frontiers in Psychology*, 17.
15. Cheng, A., Calhoun, A., & Reedy, G. (2025). Artificial intelligence-assisted academic writing: recommendations for ethical use. *Advances in Simulation*, 10(1), 22.
16. de Cassai, A., Dost, B., Augoustides, J. G. T., Azamfirei, L., Alanoğlu, Z., Azi, L. M. T. A., Calvache, J. A., Černý, V., De Hert, S. G., Eldawlatly, A., Farber, M. K., Sobreira-Fernandes, D., Fettiplace, M. R., Galante, D., Garg, R., Goldstein, H. V., Abad-Gurumeta, A., Gupta, L., Hemmings, H. C.,...Zdanowski, S. (2026). Responsible use of large language models in manuscript authorship, peer review, and editorial processes: a Delphi consensus among editors-in-chief of anaesthesia and pain medicine journals (RULE-AP). *British journal of anaesthesia*.
17. Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. B. (2023). From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, 40(2), 615–622.
18. Dianelo, R. F. B., Salac, M. B. M., Roque, A. B., Rivera, J. G., Escalona, J. M. S., De Leon, A. A. G., & Miranda, J. P. P. (2026). AI-Assisted Academic Writing in Integrated Learning Environments: A Cross-Disciplinary Scoping Review of Computational Writing Technologies. *Proceedings of the 2026 3rd International Conference on Innovation Management and Information Systems*.

19. Earp, B., Guernon, A.-S., & Mann, S. (2026a). A Substantial Human Contribution: Do We Deserve to be Authors of This Essay? *preprint ResearchGate.com*.
20. Earp, B., Guernon, A.-S., & Mann, S. (2026b). Authorship, Schmauthorship. *preprint ResearchGate*.
21. Earp, B. D., Guernon, A.-S., & Porsdam Mann, S. (2026). Thinking is not only writing. *Nature Reviews Bioengineering*, 1–2.
22. Earp, B. D., & Mann, S. P. (2026). Mapping human and AI contributions across stages of academic writing Using SLICE—a preliminary LLM-based method. *Preprint* www.researchgate.net/publication/406991607.
23. Earp, B. D., Yuan, H., Koplin, J., & Porsdam Mann, S. (2025). LLM use in scholarly writing poses a provenance problem. *Nature Machine Intelligence*, 7(12), 1889–1890.
24. Formosa, P., Bankins, S., Matulionyte, R., & Ghasemi, O. (2024). Can ChatGPT be an author? Generative AI creative writing assistance and perceptions of authorship, creatorship, responsibility, and disclosure. *AI & SOCIETY*, 40, 3405 – 3417.
25. Getto, G., Kelley, S., & Vance, B. (2026). How to Write With GenAI: A Framework for Using Generative AI to Automate Writing Tasks in Technical Communication. *Journal of Technical Writing and Communication*, 56(2), 111–144.
26. Hamdalla, M. Z. M., & Roy, S. (2026). Can we trust AI-assisted technical writing in electromagnetics? characteristic mode analysis for metasurface design as a stress test. *Frontiers in Antennas and Propagation*.
27. Hrona, N., & Semenog, O. (2026). Human-AI Academic Writing, Integrity, and Language-Communicative Culture in Teacher Education: Evidence from Ukrainian Student Texts. *Dialogica*.
28. Hu, H.-Z., Zhou, Q., & Hashim, H. (2025). Negotiating identity in the age of ChatGPT: non-native English researchers' experiences with AI-assisted academic writing. *Humanities and Social Sciences Communications*, 12.
29. Huang, L., & Liu, M. (2025). The Application of Prompt Engineering in AI-Assisted Writing: A Cross-Domain Bibliometric Literature Review (2020-July 2025). *Proceedings of the 2nd International Conference on Intelligent Computing and Data Analysis*.
30. Hurshman, C., Mann, S., Savulescu, J., & Earp, B. (2025). Authorship without writing: large language models and the "Senior Author" analogy. *preprint 10.48550/arXiv.2509.05390*.
31. Jandildinov, M. K., & Yersultanova, G. T. (2024). AI-Assisted Scholarly Writing In Education: A Scoping Review (2019–2024). *BULLBulletin Series Of Pedagogical Sciences*.
32. Jandildinov, M. K., Yersultanova, G. T., & Baissydyk, I. B. (2026). AI-Assisted Scholarly Writing in Applied Linguistics (2020–2025): A Prisma-Guided Literature Review. *Bulletin of Toraighyrov University. Pedagogics series*.
33. Jusriati, Kurniadi, W., & Nabila, A. (2026). The Effect of AI-Assisted Writing on Students' Academic Writing Skills. *English Review: Journal of English Education*, 14(2), 701–710.

34. Koo, M., & Lu, M.-C. (2026). Who Owns the Argument? A Practical Test for Author Responsibility in AI-Assisted Scholarly Writing. *Learned Publishing*, 39.
35. Li, M.-H. (2026). Generative AI-Assisted Academic Writing in Higher Education EFL Contexts: A Critical Review of Learning Outcomes, Self-Regulated Learning, and Ethical Governance. *Journal of English Language Teaching and Applied Linguistics*.
36. Lu, C., Lu, C., Lange, R. T., Yamada, Y., Hu, S., Foerster, J., Ha, D., & Clune, J. (2026). Towards end-to-end automation of AI research. *Nature*, 651(8107), 914–919.
37. Lund, B. D., & Naheem, K. T. (2023). Can ChatGPT be an author? A study of artificial intelligence authorship policies in top academic journals. *Learned Publishing*, 37.
38. Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, 74(5), 570–581.
39. Meng, Q., Xie, Y.-N., & Lu, D. (2026). A systematic review of AI-assisted academic writing: Tools, measures, and effects. *Journal of English for Academic Purposes*.
40. Metcalf, T. (2026). Is AI-Produced Humanities Scholarship a Case of Research Misconduct? *Journal of Academic Ethics*, 24.
41. Nazarovets, S., & Teixeira da Silva, J. A. (2025). ChatGPT as an “author”: Bibliometric analysis to assess the validity of authorship. *Accountability in Research*, 32(7), 1148–1158.
<https://doi.org/10.1080/08989621.2024.2345713>
42. Neshkovska, S. (2025). The Merits and Risks of AI-Assisted Academic Writing: Insights from Current Research. *ELOPE: English Language Overseas Perspectives and Enquiries*.
43. Nguyen, A., Hong, Y., Dang, B., & Huang, X.-S. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education*, 49, 847 – 864.
44. Paltridge, B. (2026). AI and scholarly writing: Ethics, authorship and ownership. *Journal of Second Language Writing*.
45. Pan, H.-C., Araujo Oliveira, E., & Ferreira Mello, R. (2025). Exploring Human-AI Collaboration in Educational Contexts: Insights from Writing Analytics and Authorship Attribution. *Proceedings of the 15th International Learning Analytics and Knowledge Conference*.
46. Pividori, M. D., & Greene, C. S. (2024). A publishing infrastructure for Artificial Intelligence (AI)-assisted academic authoring. *Journal of the American Medical Informatics Association : JAMIA*, 31, 2103 – 2113.
47. Porsdam Mann, S., Earp, B. D., Nyholm, S., Danaher, J., Møller, N., Bowman-Smart, H., Hatherley, J., Koplin, J., Plozza, M., & Rodger, D. (2023). Generative AI entails a credit–blame asymmetry. *Nature Machine Intelligence*, 5(5), 472–475.
48. Porsdam Mann, S., Vazirani, A. A., Aboy, M., Earp, B. D., Minssen, T., Cohen, I. G., & Savulescu, J. (2024). Guidelines for ethical use and acknowledgement of large language models in academic writing. *Nature Machine Intelligence*, 6(1), 1272 – 1274.

49. Pratama, R. D., Widiati, U., Basthomi, Y., El Khoiri, N., Muniroh, S., & Omolu, F. A. (2026). "From Textual Originality to Cognitive Transparency": Reframing Ethics and Authenticity of Scientific Writing in the Age of Artificial Intelligence through Human-AI Collaborative Writing. *Journal of Academic Ethics*, 24.
50. Rahmat, N. H., Khairuddin, Z., Aripin, N., & Mokhtar, M. I. (2026). Scaffolding Academic Writing Using AI Assistance: A Systematic Literature Review. *International Journal of Research and Innovation in Social Science*.
51. Roa, A. A. P., & Halim, S. (2024). The Impact of AI-Powered Software on Second Language (L2) Writing: A Systematic Literature Review. *Research and Innovation in Applied Linguistics-Electronic Journal*.
52. Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Can artificial intelligence help for scientific writing? *Critical care*, 27(1), 75.
53. Shen, X., & Tao, Y.-H. (2024). Metacognitive Strategies, AI-based Writing Self-efficacy and Writing Anxiety in AI-assisted Writing Contexts: A Structural Equation Modeling Analysis. *International Journal of TESOL Studies*.
54. Silva, J., Gouveia, R. P., Zielinski, K. M. C., Oliveira, M. C. F., Amancio, D. R., Bruno, O. M., & Oliveira, O. N. (2025). AI-Assisted Tools for Scientific Review Writing: Opportunities and Cautions. *ACS Applied Materials & Interfaces*, 17, 47795 – 47805.
55. Slimi, Z. (2026). A systematic critical review of generative AI's impact on authorship, pedagogy, and integrity (2023–2025). *Frontiers in Education*.
56. Sudirman, A., Kristanto, T., Vitalya Gemilang, A., & Rais, B. (2026). Harnessing AI-assisted Writing in Educational Institutions: A Systematic Literature Review. *IJORER : International Journal of Recent Educational Research*.
57. Tang, W.-Q., Wang, Z., & Yan, G.-F. (2026). What Are We Missing in AI-Supported L2 Writing Research? A Systematic Review From 2014 to 2024. *Journal of Computer Assisted Learning*.
58. Uebel, M., & Hamamra, B. T. (2026). Authorship After Generative AI: Distributed Creativity and Relational Responsibility. *Philosophy & Technology*, 39.
59. Urzúa, C. A. C., Ranjan, R., Saavedra, E. E. M., Badilla-Quintana, M. G., Lepe-Martínez, N., & Philominraj, A. (2025). Effects of AI-Assisted Feedback via Generative Chat on Academic Writing in Higher Education Students: A Systematic Review of the Literature. *Education Sciences*.
60. van Woudenberg, R., Ranalli, C., & Bracker, D. (2024). Authorship and ChatGPT: a Conservative View. *Philosophy & Technology*, 37(1), 34.
61. Watson, S., Brezovec, E., & Romic, J. (2025). The role of generative AI in academic and scientific authorship: an autopoietic perspective. *AI & SOCIETY*, 40, 3225 – 3235.
62. Xu, J., Lu, X., & Abdi, N. M. (2026). Re-centering the role of humans in AI-assisted writing scoring: A scoping review from a critical perspective. *Assessing Writing*.
63. Yoo, J. (2025). Defining the Boundaries of AI Use in Scientific Writing: A Comparative Review of Editorial Policies. *Journal of Korean Medical Science*, 40.

64. Zadakhanova, A. A., Janserkeyeva, E., Montgomery, D., & Baissydyk, I. (2025). AI-Assisted Scholarly Writing In Applied Linguistics and Language Education In Kazakhstan: A Scoping Review. *Journal of Educational Sciences*, 150–168.

Figures

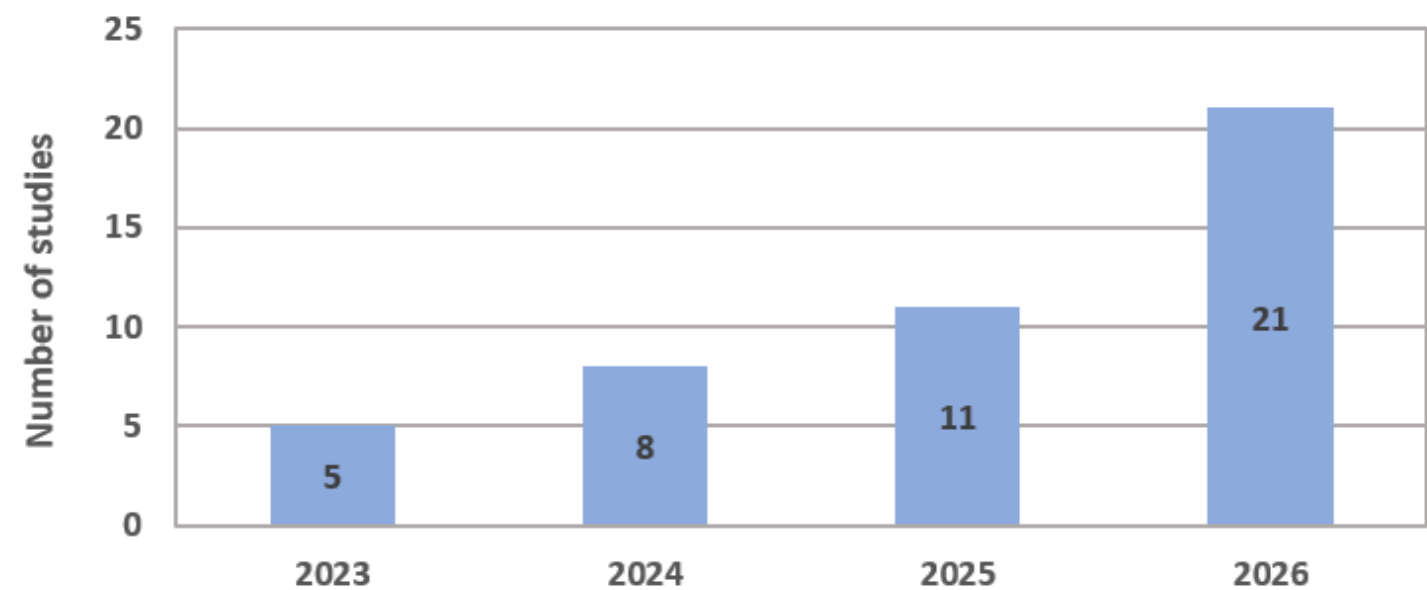


Figure 1

Number of included primary studies by publication year

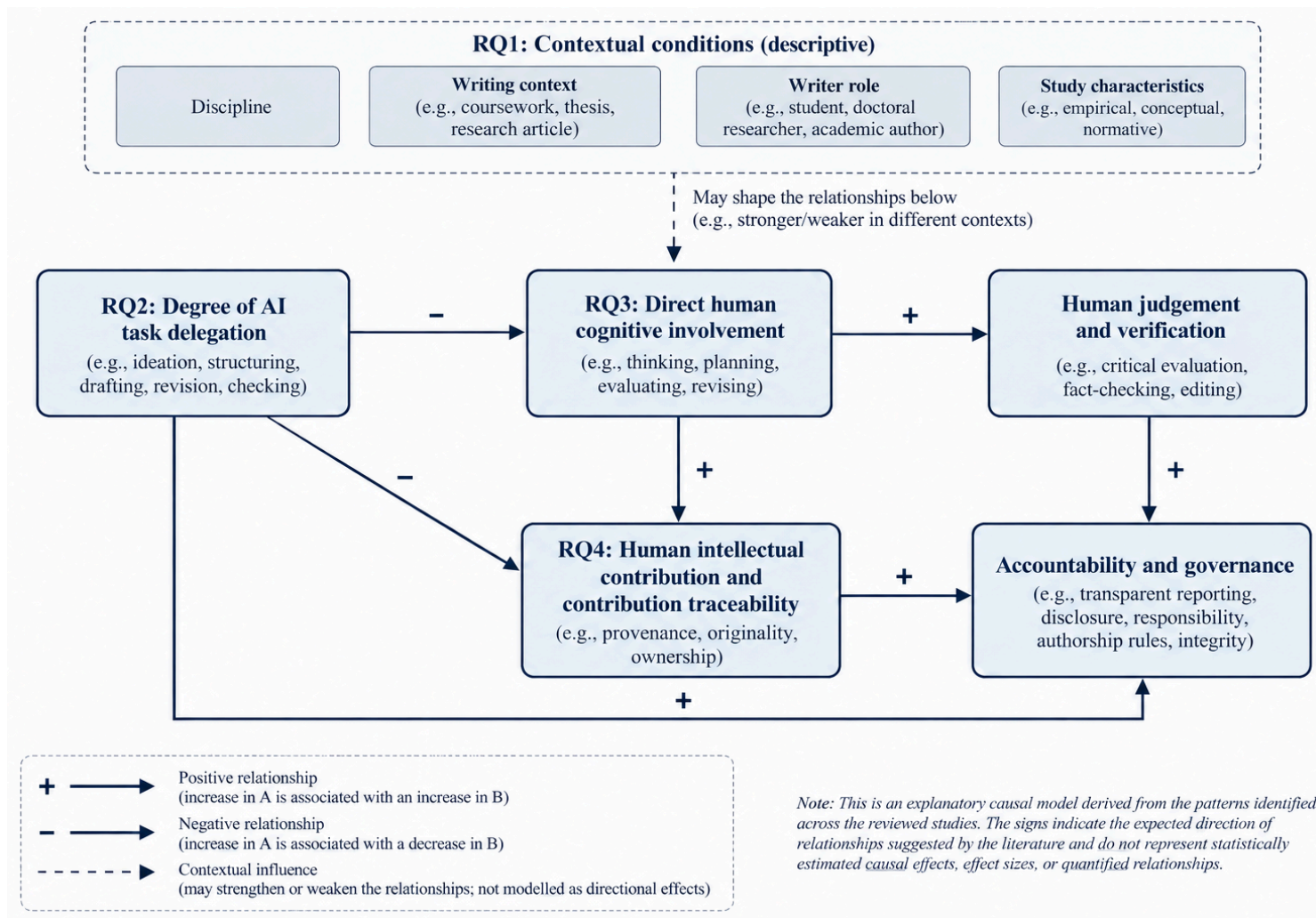


Figure 2

A literature-derived explanatory causal model of human–AI academic co-writing.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial.docx](#)
- [Appendixs.docx](#)