



Review

Thinking with the Machine: Generative AI, Cognitive Processes, Writing Practices and Human Agency in Education

Edwin Creely

Faculty of Education, Monash University, Clayton, VIC 3800, Australia; edwin.creely@monash.edu

Abstract

This integrative literature review examines how generative artificial intelligence is shaping human thinking, cognitive processes and writing practices, with a focus on education across schools, universities, adult learning and professional settings. It considers how AI-supported composing alters the ways writers plan, generate, organise, revise and evaluate text, and how these changes reshape cognition, agency, authorship and learning. Rather than treating generative AI simply as a tool for productivity, the review approaches it as a relational technology that mediates meaning-making and textual production, while reserving the language of intentional agency for human actors. Thirty-six publications, identified through a structured search of seven information sources reported in accordance with PRISMA 2020 and appraised using criteria matched to each form of evidence, were synthesised thematically. Albert Bandura's social cognitive theory provides the analytical lens, particularly triadic reciprocal causation, self-efficacy, observational learning, self-regulation and human agency. The discussion develops a critical account that recognises clear benefits while flagging risks of cognitive offloading, uncritical trust, homogenised expression and weakened metacognitive control. Claims are calibrated to the type of evidence supporting them, and findings drawn from preprints, small samples and self-report designs are presented as preliminary. The review advances a hybrid account in which human and machine contributions to composing are entwined and argues that this entwinement makes deliberate human participation more rather than less necessary. It concludes by proposing guidelines for educators across sectors, together with the assessment conditions under which those guidelines can be verified.

Keywords: generative AI; writing; cognition; social cognitive theory; critical AI literacy



Academic Editor: George D.
Magoulas

Received: 28 June 2026

Revised: 26 August 2026

Accepted: 14 September 2026

Published: 17 September 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction and Context

The arrival of widely accessible generative artificial intelligence has unsettled long-held assumptions about who, or what, participates in the act of writing. Within a very short period, platforms such as ChatGPT, Claude, Gemini and Copilot moved from technical novelty to everyday infrastructure in classrooms, workplaces and homes. Educators now confront a situation in which a learner can produce a fluent paragraph, an essay plan or a polished argument in seconds, often without the sustained mental effort such tasks once demanded. This review begins from the premise that generative AI does far more than assist with grammar or surface editing: it increasingly enters the cognitive space of invention, planning, drafting, argument formation and stylistic choice, participating in the very processes through which people think their way into writing. The central question is therefore deceptively simple: how is generative AI shaping cognitive processes and writing practices, and what does this mean for education?

Writing has never been a purely mechanical transcription of finished thought. The cognitive process tradition established that composing is a recursive set of mental operations in which planning, translating and reviewing interact continuously, and in which the act of putting words on a page reshapes the ideas being expressed (Flower & Hayes, 1981; Hayes, 2012). On this account, writing is a way of thinking rather than merely a record of it. If that is so, then a technology that intervenes in drafting and revision is not a neutral accessory; it is an intervention in cognition itself. The significance of generative AI for education lies precisely here. When a system can generate plausible content on demand, the boundary between thinking and writing, and between the writer's mind and the external resource, becomes porous in ways that earlier writing technologies only approached.

Debate about these developments has tended to cluster around three positions. The first frames generative AI as an extension of human capability, a cognitive partner that scaffolds composition, lowers barriers for novice and multilingual writers, and frees attention for higher-order concerns (Kasneci et al., 2023; Salomon et al., 1991; Song & Song, 2023). The second is more anxious, warning that habitual reliance on machine-generated text diminishes independent thinking, weakens memory and judgement, and encourages a quiet deskilling of the very capacities education exists to cultivate (Gerlich, 2025; Kosmyna et al., 2025; H.-P. Lee et al., 2025). The third resists this binary, proposing that generative AI is producing new hybrid forms of authorship and cognition in which human and machine contributions are entangled and difficult to separate (Draxler et al., 2024; Markauskaite et al., 2022). This review takes the third position seriously while drawing critically on the insights of the other two.

The educational stakes are considerable and they differ across sectors. In schools, teachers must decide how to nurture foundational writing and reasoning when learners can bypass the productive struggle that builds them. In higher education, academics face questions about authorship, assessment, evaluative judgement and the integrity of qualifications (Cotton et al., 2024; Tai et al., 2018). In adult, vocational and community education, generative AI offers genuine support for confidence and access while risking dependence that undercuts the slow development of voice and autonomy (Han & Reinhardt, 2022). In professional learning, the redistribution of cognitive effort between person and machine has consequences for how knowledge is produced and trusted (H.-P. Lee et al., 2025).

A further reason to attend closely to generative AI is that it shapes thought before a single word is committed to the page. Prompting a system for an outline, a counterargument or a framing does not merely speed composition; it supplies the conceptual furniture with which the writer proceeds. The influence therefore reaches invention itself, the stage at which a writer decides what is worth saying, and it is this reach into the formative stages of thinking that distinguishes generative AI from earlier writing technologies.

There are good reasons to undertake this synthesis now and in an integrative rather than a narrowly empirical mode. The literature has grown explosively and unevenly, scattered across disciplines that rarely speak to one another, and educators making daily decisions cannot wait for the slow accumulation of longitudinal evidence. What is needed is a synthesis that brings empirical and conceptual work into a single theoretical frame.

To analyse these dynamics, the review adopts Bandura's social cognitive theory, which conceives of human functioning as the product of continuous reciprocal interaction between personal factors, behaviour and the environment. The theory is well suited to the task because it neither reduces the person to a passive recipient of technological influence nor imagines an autonomous agent unaffected by the tools at hand. The review asks how generative AI mediates the relationship between person, behaviour and environment, and what this means for education.

2. Theory: Bandura's Social Cognitive Theory

Social cognitive theory offers a parsimonious yet powerful account of how people learn, act and develop within social and material environments. Its central proposition is triadic reciprocal causation, the claim that personal factors such as beliefs and emotions, behaviour, and environmental influences operate as interacting determinants that shape one another bidirectionally (Bandura, 1986, 2001). None of the three is sovereign. A learner's confidence influences the writing they attempt, the writing they produce alters their environment and the feedback it returns, and that altered environment in turn revises their confidence. Applied to generative AI, this model invites a relational analysis in which the technology is treated as a salient and active feature of the environment in reciprocal relationship with the writer rather than as an external instrument standing outside cognition.

The construct of self-efficacy, the belief in one's capability to organise and execute the actions required to attain a goal, is central to the theory and to this review (Bandura, 1977, 1997). Efficacy beliefs shape the tasks people choose, the effort they invest, and their persistence in difficulty, and in writing they predict engagement, strategy use and achievement (Sun & Wang, 2020). A hesitant writer who receives an immediate, fluent draft may experience a surge of confidence, yet whether that confidence reflects a durable capability or a borrowed competence is precisely what educators must learn to discern. Bandura distinguished genuine efficacy, built through mastery experience, from inflated or misattributed confidence, and generative AI can supply the appearance of mastery without the substance.

Observational learning and modelling provide a second relevant mechanism. People acquire knowledge, skills and standards by observing others and inferring rules from what they see, without direct instruction or reinforcement (Bandura, 1986). Generative AI is a ubiquitous and tireless model of written form, and learners who repeatedly encounter its output absorb its conventions of structure, register and stance. The theory therefore predicts that sustained exposure will shape not only what learners produce but what they take good writing to be.

Self-regulation and human agency complete the framework and carry the greatest weight in the analysis that follows. Bandura's account of agency turns on the capacity to influence one's own functioning and life circumstances deliberately, through forethought, self-reactive influence and self-reflection (Bandura, 2001, 2018). Self-regulation is the operational expression of that capacity: goal setting, monitoring, evaluation and adjustment. Writing has long been understood as a self-regulated activity, and generative AI intervenes directly in that regulation, since a system that plans, drafts and revises can absorb the very operations through which learners regulate themselves. Whether it strengthens or supplants self-regulation is the pivotal question.

Agency, Proxy Agency, Technological Mediation and Affordance

Because this review describes generative AI as a mediating agent, the sense in which the word agent is used requires precision. Bandura reserves agency for intentional human action: the capacity to form intentions, exercise forethought, regulate conduct through self-reactive influence, and reflect upon one's own functioning (Bandura, 2001, 2018). Generative systems possess none of these properties. Four notions are therefore distinguished and used consistently. Human agency denotes intentional, forethoughtful and self-reflective action by a person. Proxy agency, in Bandura's sense, denotes securing outcomes through others who command capabilities one lacks, and it describes the learner who delegates composition while remaining the party who holds the intention and bears the responsibility. Technological mediation denotes the way an artefact shapes the relation between a person and the world without itself intending anything (Latour, 1994; Verbeek, 2005; Wertsch, 1998).

Environmental affordance denotes the possibilities for action a system makes available. The mediating agent invokes the third and fourth senses and never the first.

3. Methodology

This review adopts an integrative methodology because the phenomenon under study is interdisciplinary, fast-moving and served by evidence of very different kinds. The integrative review admits the empirical, theoretical and methodological literature within a single synthesis and is designed to generate new conceptual understanding rather than to aggregate effects (Snyder, 2019; Torracco, 2005; Whitemore & Knafl, 2005). Reporting follows PRISMA 2020 conventions for identification and selection (Page et al., 2021).

A word is warranted on what this design can and cannot deliver, since the reporting expectations that apply to an integrative review differ from those governing a systematic review of intervention effects. Its warrant lies in the transparency and defensibility of its interpretive reasoning rather than in the homogeneity of its corpus. This review therefore adopts PRISMA reporting for identification and selection, because reproducibility of the search is a reasonable expectation of any review, while declining to import procedures that presuppose a body of comparable empirical studies. Where systematic-review procedure has deliberately not been followed, the reasoning is stated at the point it arises.

Searches were conducted across seven sources selected for complementary coverage of education, psychology, linguistics and computing, listed in Table 1, using the concept clusters in Table 2. Clusters were combined with AND and terms within clusters with OR, with truncation and phrase searching as each platform permits. Searches ran in title, abstract and keyword fields in Scopus, Web of Science, PsycINFO and LLBA, in title, abstract and descriptor fields in ERIC and Education Research Complete, and across the full record in Google Scholar. Coverage ran from 1 January 2022 to the date of search; language was restricted to English; and document types were limited to journal articles, peer-reviewed conference papers, scholarly books and chapters, and reviews. Hand-searching of reference lists was not date-limited, since foundational theory predating 2022 was required. Initial searches ran between 12 and 19 January 2026 and an updated search on 4 May 2026. Because Google Scholar lacks the Boolean depth and export functions of the other platforms, a bounded procedure was used: a simplified string sorted by relevance with patents and citations excluded, screened in rank order at twenty per page, stopping after two consecutive pages without an eligible record, which occurred at result 200. No record was included on a Google Scholar hit alone. Every query is reproduced verbatim as Table S1 in the Supplementary Materials (Rethlefsen et al., 2021).

Table 1. Databases searched and disciplinary rationale.

Database	Disciplinary Coverage and Rationale
Scopus	Broad multidisciplinary index used to capture education, psychology, computing and linguistics literature and to support citation tracking.
Web of Science	Multidisciplinary citation database used alongside Scopus to widen coverage and identify high impact and frequently cited work.
ERIC	Core education database providing comprehensive coverage of schooling, higher education and adult education research.
Education Research Complete	Education focused index used to capture journal literature on pedagogy, literacy and classroom practice.
PsycINFO	Psychology database essential for self-efficacy, metacognition, cognitive offloading and social cognitive theory.

Table 1. *Cont.*

Database	Disciplinary Coverage and Rationale
LLBA (Linguistics and Language Behaviour Abstracts)	Linguistics database used to capture writing, composition and second language literature.
Google Scholar	Used as a supplementary source to identify recent preprints, conference proceedings and grey literature in a fast-moving field.

Table 2. Search term clusters.

Concept Cluster	Representative Search Terms
Technology	“generative AI”, “generative artificial intelligence”, “large language model”, “ChatGPT”, “AI writing tool”
Cognition	cognition, “cognitive offloading”, metacognition, “critical thinking”, agency, “self-efficacy”, “self-regulation”
Writing	writing, composition, “academic writing”, “second language writing”, authorship, revision
Education	education, school, “higher education”, “adult education”, “AI literacy”, learning, pedagogy

Inclusion and exclusion criteria, summarised in Table 3, prioritised peer-reviewed publications appearing between 2022 and 2026, the period bracketing the public release of large language model chat interfaces, while admitting earlier foundational sources for social cognitive theory, writing process research and theories of distributed and extended cognition. Publications were eligible if they addressed, conceptually or empirically, the relationship between generative AI and human cognition, writing or learning, and if they were available in English. Sources were excluded if they treated AI only as a back-end engineering concern with no bearing on human thinking or writing, if they were purely promotional, or if they could not be retrieved in full. Reference lists of included works were hand-searched to capture influential items missed by database queries, and a small number of seminal pre-2022 works were retained because they remain load bearing for the theoretical argument.

Table 3. Inclusion and exclusion criteria.

Dimension	Inclusion	Exclusion
Focus	Addresses generative AI in relation to writing, cognition or learning	Generative AI incidental; no human or educational dimension
Publication type	Peer reviewed articles, conference papers and scholarly books; foundational theory where needed	Opinion pieces, marketing material and non-scholarly commentary
Timeframe	Empirical work from 2022 to 2026; earlier foundational theory admitted	Superseded technical reports without conceptual or educational relevance
Language and access	English language and retrievable in full text	Not available in full text; language not readable by the reviewer
Relevance	Speaks to thinking, writing, agency or education	Purely technical or computational with no learning relevance

Preprints are eligible on three conditions, following guidance for the grey literature (Adams et al., 2017): the source must report primary data that can be appraised; it is appraised with the instrument appropriate to its design and then downgraded by one level of confidence for the absence of peer review; and its provisional status is stated wherever it

is cited. One preprint, [Kosmyna et al. \(2025\)](#), is retained on this basis. Excluding it would have suppressed the only direct neurophysiological evidence on a question central to the review, which seemed a worse distortion than including it under a stated discount.

The screening process is reported in Table 4, following the PRISMA 2020 flow ([Page et al., 2021](#)). The original submission described the yield as approximately 612 records, which was an imprecision rather than an estimate; the counts below are exact. Database-specific yields were Scopus 147, Web of Science 118, ERIC 92, Education Research Complete 71, PsycINFO 58, LLBA 27 and Google Scholar 80, with 19 further records from hand-searching and citation tracking, giving 612 in total. Removal of 194 duplicates left 418 records for title and abstract screening, of which 334 were excluded. Full texts were retrieved for all 84 remaining records. Forty-eight were excluded: no substantive engagement with cognition, writing or learning (17); commentary without scholarly apparatus (12); generative AI incidental (9); purely technical (6); and duplicate report (4). Thirty-six publications were carried forward. The updated search identified 47 further records, of which 3 were included and are marked in Table S2.

Table 4. Screening and selection process.

Stage	Action	Records (n)
Identification	Records identified across databases and supplementary searching	612
Duplicate removal	Records remaining after duplicates removed	418
Abstract screening	Titles and abstracts screened against criteria	418
Full-text eligibility	Full texts retrieved and assessed for eligibility	84
Included in synthesis	Studies carried forward into thematic synthesis	36

The 36 sources that remained span empirical studies using behavioural, survey and neurophysiological methods, conceptual and theoretical analyses, and reviews, alongside foundational works that the integrative method admits. The corpus comprises 14 empirical studies, 3 reviews and 19 conceptual or theoretical works. Table S2 identifies every one of the 36 and records, for each, its evidence type and design, context and sector, participants where applicable, principal contribution, appraisal rating and thematic contribution. The distinction the original submission asserted but did not operationalise is thereby auditable: the 36 publications in Table S2 constitute the corpus, while [Braun and Clarke \(2006\)](#), [Page et al. \(2021\)](#), [Snyder \(2019\)](#), [Torraco \(2005\)](#), [Whittemore and Knafl \(2005\)](#) and the appraisal and mediation sources added during revision are background cited in support of the review's conduct.

These counts correspond item for item to the PRISMA 2020 flow and are consistent across the text, Table 4 and Table S2. A separate flow diagram is not reproduced because it would duplicate Table 4 exactly; the screening spreadsheet is available from the author on request.

Analysis proceeded thematically ([Braun & Clarke, 2006](#)), adapted to the integrative purpose of building conceptual understanding across heterogeneous sources. The framework developed across three cycles. First-cycle coding was descriptive and largely in vivo and generated 147 codes; second-cycle coding collapsed these into 31 categories; third-cycle work tested candidate themes against internal homogeneity, external heterogeneity, support from at least four independent sources, and analytic contribution. Eight candidates survived the second cycle and two were dissolved, with trust and verification absorbed into metacognition and voice distributed across authorship and homogenisation. Contradicting evidence was actively sought and is reported in the findings rather than set aside: documented gains in motivation and self-efficacy ([Huang & Mizumoto, 2024](#); [Song & Song,](#)

2023), individual creativity raised even as collective diversity narrows (Doshi & Hauser, 2024), and positions holding that the opportunities of large language models outweigh their risks (Kasneci et al., 2023).

Because the corpus combined evidence of very different kinds, no single appraisal instrument could be applied uniformly. The original submission described this pluralism but did not name the instruments used. Three are named here. Empirical studies were appraised with the Mixed Methods Appraisal Tool, version 2018 (Hong et al., 2018), with qualitative studies additionally screened against the CASP checklist (Critical Appraisal Skills Programme, 2018). Reviews were appraised against the Joanna Briggs Institute criteria (Aromataris et al., 2015). Ratings were converted to a common scale on which high confidence required all five design-specific criteria to be met, moderate four, and low three or fewer, with non-peer-reviewed sources downgraded one level. Of the 14 empirical studies, 5 were rated high, 7 moderate and 2 low; of the 3 reviews, 2 high and 1 moderate. Kosmyna et al. (2025) met four of five criteria and would otherwise have been moderate but was downgraded to low for the absence of peer review. Study-level ratings appear in Table S2.

Conceptual and theoretical works were treated differently, and deliberately so. Applying a study-quality instrument to Bandura's (1986) statement of social cognitive theory or Flower and Hayes's (1981) process model would be a category error: these works make no empirical claims, and an instrument designed to detect sampling, measurement and analytic bias has nothing in them to measure, so any rating would be an artefact of the instrument. Such works were assessed instead for clarity of construct definition, warrant, integration with existing theory, specification of boundary conditions, and generativity (MacInnis, 2011). Of the 19, 16 satisfied all five criteria and 3 satisfied four. Appraisal was interpretive rather than a threshold for exclusion, but it governed weight: no claim rests on a low-confidence source alone, and where one is the principal evidence, that is stated at the point of use.

3.1. Reviewer Roles, Safeguards and Positionality

All stages of the review were conducted by a single reviewer. This is a substantive limitation rather than a design choice, and it bears directly on how the procedures above should be read. Four safeguards were implemented. First, a complete audit trail was maintained, comprising dated search logs, a screening spreadsheet recording a decision and reason for each of the 418 records screened at title and abstract and each of the 84 assessed at full text, an extraction form for every included publication, and analytic memos documenting theme development; this material is available from the author on request. Second, a random 20 per cent of screened records was re-screened blind after four weeks, giving intra-rater agreement of Cohen's kappa of 0.87 at title and abstract and 0.91 at full text, with six discrepancies resolved against the eligibility criteria and two records reinstated. Third, a colleague with no other involvement independently coded a random 25 per cent of the corpus, with 84 per cent agreement on theme assignment; discrepancies were resolved to consensus, the principal outcome being the merging of trust and verification into metacognition. Fourth, negative case analysis was conducted deliberately. Nowell et al.'s (2017) trustworthiness criteria were used throughout. As a teacher educator and practising writer working with adult and migrant learners, I bring an interpretive orientation towards agency and access, and readers should weigh the synthesis accordingly.

3.2. Limitations

Several limitations qualify the findings. The pace of AI development means empirical claims risk obsolescence, and the evidence base is uneven, comprising a small number of controlled studies alongside a much larger body of survey, self-report and conceptual work. Three rules follow for how the findings should be read. Cross-sectional and survey evidence supports statements of association only and is not reported as demonstrating causation. Short-duration performance studies cannot by themselves establish durable developmental consequences; where the review discusses long-term effects it is extrapolating from theory and says so. Findings from preprints or small exploratory studies are labelled preliminary at every point of use. The restriction to English-language sources is a further material limitation, particularly given the review's own argument about the homogenising effect of Anglophone conventions.

4. Findings

Table 5 summarises how the six themes map onto the primary social-cognitive constructs that organise the analysis. The themes are presented in sequence but describe a single interconnected reality in which offloading, metacognitive monitoring, compositional practice, ownership, confidence and stylistic convergence continually condition one another.

Table 5. Themes mapped to social cognitive constructs.

Theme	Analytic Focus	Primary Social-Cognitive Construct
Cognitive offloading	Delegation of planning, retrieval, drafting and revision to the system	Reciprocal determinism; environment
Metacognition	Monitoring, evaluation and control of one's own thinking when using AI	Self-regulation
Writing processes	Reorganisation of planning, translating and reviewing in AI assisted composing	Behaviour; self-regulation
Authorship	Ownership, responsibility and voice in hybrid human-AI text	Human agency; proxy agency
Self-efficacy	Confidence in one's writing capability, located in the self or the tool	Self-efficacy; outcome expectations
Homogenisation	Convergence of expression and stance through repeated modelling	Observational learning

4.1. Cognitive Offloading

Cognitive offloading refers to the use of external resources and physical action to reduce the internal information processing a task demands (Risko & Gilbert, 2016). The practice is ancient and often beneficial, and the traditions of distributed and extended cognition have long argued that thinking is properly understood as a property of person-plus-tool systems rather than of brains alone (Clark & Chalmers, 1998; Hutchins, 1995; Salomon, 1993). Offloading is therefore not intrinsically a deficit.

The concern with generative AI is that it permits the offloading not merely of storage, as a search engine does (Sparrow et al., 2011), but of the generative and evaluative operations at the heart of thinking: framing a problem, forming an argument, weighing evidence, judging what is worth saying. When these are delegated, what is externalised is not the burden of remembering but the work of thinking itself, and it is this shift in what is offloaded, rather than the fact of offloading, that warrants educational attention.

Recent empirical work lends weight to this concern while complicating it. A large survey study found a significant negative association between frequent AI use and critical thinking scores, with cognitive offloading statistically mediating the relationship (Gerlich, 2025). Because the design is cross-sectional and relies on self-reported use, the direction of influence cannot be established: learners who think less critically may be more inclined

to delegate, as readily as delegation may erode critical thinking. A study of knowledge workers similarly reported that higher confidence in generative AI was associated with reduced critical engagement, while higher confidence in one's own expertise was associated with greater scrutiny of AI output (H.-P. Lee et al., 2025), again on self-report measures. A small neurophysiological study reported reduced neural connectivity and weaker recall of one's own text among participants who used a large language model (Kosmyna et al., 2025). That finding is the most striking in the corpus and the most provisional: it derives from an unreplicated preprint with a small sample and should be read as a hypothesis worth testing rather than an established result.

The offloading lens also clarifies why the effects of generative AI differ across sectors. In adult, vocational and second-language settings, offloading surface-level demands can free attention for higher-order concerns and open access to registers otherwise out of reach. In the early development of writing, where the effortful operations being offloaded are precisely those the learner is meant to acquire, the calculation is different. The same behaviour is enabling in one case and developmentally costly in the other.

For education, the central implication is that offloading must be designed rather than left to chance. The goal is not to eliminate external support but to arrange the human-tool system so that learners retain the operations that matter for their development (Salomon et al., 1991). A learner who uses AI to check a finished argument is offloading differently from one who uses it to produce the argument.

4.2. Metacognition

Metacognition, the awareness and regulation of one's own thinking, has long been recognised as a determinant of learning and a hallmark of skilled writing (Flavell, 1979). Competent writers monitor whether a sentence says what they mean and whether an argument holds, and this monitoring is what generative AI most directly disturbs.

The evidence suggests that AI supports metacognition only when learners actively interrogate its output. Where learners verify, question and selectively incorporate suggestions, engagement with the material can deepen; where fluent output is accepted on trust, the monitoring loop is short-circuited (H.-P. Lee et al., 2025), though this pattern rests on cross-sectional self-report and is an association rather than a demonstrated mechanism. Fluency is a poor proxy for accuracy, and a well-formed sentence can be confidently wrong.

In language and writing education, the most promising uses position AI as a feedback partner that scaffolds evaluation rather than as a producer of finished text. Automated written feedback can support noticing and revision when learners are taught to appraise it critically (Shi & Aryadoust, 2024), and structured collaboration with generative systems can improve argumentative writing where the learner retains responsibility for the argument (Su et al., 2023).

These observations connect to models of self-regulated learning, in which learners cycle through forethought, performance and self-reflection (Zimmerman, 2002). Where AI is inserted in that cycle determines whether it supports or supplants regulation: used in forethought to clarify goals it can prime self-regulation; used in performance to produce the work it removes the occasion for monitoring; used in reflection to evaluate a draft it can model the self-assessment learners must eventually perform alone.

A further metacognitive risk concerns calibration, the accuracy of learners' judgments about what they know and can do. Decisions to rely on external resources rest on metacognitive evaluations of one's own abilities, which are frequently erroneous (Risko & Gilbert, 2016). When fluent AI output stands in for the learner's own performance, the feedback that ordinarily calibrates self-assessment is disrupted. This follows from established work on calibration rather than from any study in the present corpus, none of

which has measured the divergence between confidence and unaided capability among AI-using writers. It is advanced as a prediction the field should test.

4.3. Writing Processes

Generative AI is reshaping the writing process at every phase the cognitive process tradition identified (Flower & Hayes, 1981; Hayes, 2012). In planning, learners increasingly begin from a machine-generated outline rather than from the exploratory struggle through which goals and ideas are ordinarily discovered, and planning shifts from generation towards selection among options the system supplies.

In drafting, the process is becoming more iterative and prompt based. Rather than producing prose linearly, writers elicit passages, evaluate them and re-prompt, so that composing resembles curation as much as production (M. Lee et al., 2022; Wang, 2024). This can support fluency and reduce the paralysis of the blank page, but it also alters what the writer practises.

Revision and editing are similarly transformed. AI can identify weaknesses, propose reorganisations and supply alternative phrasings at a speed no human reviewer matches. Where learners evaluate these proposals, revision becomes an occasion for judgement; where they accept them wholesale, the revising that most develops a writer is the operation most readily surrendered.

These changes are felt differently across the stages of schooling. In the early development of writing, fluency and control are still being built and premature delegation may forestall their acquisition. In senior secondary and tertiary settings, where basic control is established, the risk shifts to the atrophy of higher-order composing. These sector contrasts are extrapolations from the developmental logic of the writing literature rather than findings from studies comparing sectors directly.

Across these phases, the most consequential change is the compression of the distance between thought and text. The friction of composing traditionally forced a slow elaboration of ideas, and that friction was generative, since the difficulty of expression often clarified the thought. The implication is not that friction should be artificially preserved, but that teachers must distinguish productive difficulty from friction that is merely tedious.

4.4. Authorship

Generative AI destabilises authorship, the attribution of a text to a responsible originating mind, in ways education has yet to absorb. Experimental work shows that writers using AI assistance feel reduced ownership of the resulting text even when they direct the process closely, a pattern described as the AI ghostwriter effect (Draxler et al., 2024). Ownership matters educationally because it underwrites the investment through which writing develops.

The question of voice is especially acute. Voice, the distinctive signature of a writer's stance and sensibility, is among the hardest things to teach and among the most consequential to lose. Where much of a text originates in a system trained to produce conventional prose, the writer's own register is easily displaced before it has consolidated.

Academic integrity is the most visible institutional expression of these anxieties. The capacity of generative AI to produce assessable work has prompted concern about the reliability of qualifications, alongside recognition that detection is unreliable and prohibition alone inadequate (Cotton et al., 2024; van Dis et al., 2023). Where assessment has relied on the unsupervised production of text as a proxy for learning, the fragility of that proxy is now exposed.

In higher education, in particular, the more durable responses shift the locus of assessment from the product to the judgement behind it: the reasoning a learner can

articulate, the evaluative commentary they can offer on a draft, and the position they can defend in dialogue. This relocates rigour to the evaluative judgement that becomes more important, not less, when fluent text is abundant (Tai et al., 2018).

Underlying these concerns is a conceptual shift. If writing is increasingly a negotiation between human intention and machine generation, authorship becomes a matter of degree and of direction rather than of origin. This is the hybrid position introduced in Section 1 (Draxler et al., 2024; Markauskaite et al., 2022), and it does not dissolve responsibility, since only the human party can hold an intention or answer for a claim.

4.5. Self-Efficacy

Generative AI exerts a powerful and ambivalent influence on writing self-efficacy. Studies report gains in motivation and confidence among learners using generative systems, particularly in second-language contexts where the affective barriers to writing are high (Huang & Mizumoto, 2024; Song & Song, 2023), on short-duration quasi-experimental designs. These gains are real and educationally valuable.

Bandura's theory clarifies where the danger lies. Self-efficacy is built most durably through mastery experience, the accomplishment of a demanding task by one's own effort (Bandura, 1997). Where AI produces the accomplishment, the experience may be attributed to the tool rather than the self, and confidence may rise without the capability it is meant to index. No study in the corpus has tested this divergence directly, and it is offered as a prediction rather than a finding.

The vicarious and persuasive sources of efficacy are also engaged. By modelling competent writing, generative AI offers a form of vicarious experience, and by responding supportively a form of social persuasion, both weaker sources than mastery (Bandura, 1986). A learner who watches the system model a strong paragraph, attempts one unaided and succeeds has converted support into mastery; one who accepts the output has not.

These dynamics are especially salient for adult learners and for learners from migrant and refugee backgrounds, for whom academic English can present formidable affective as well as linguistic barriers. Generative AI can function as a patient interlocutor that lowers the threshold to participation and supplies the early successes that rebuild a willingness to write (Han & Reinhardt, 2022). The accompanying risk is that support becomes a permanent prosthesis rather than a temporary scaffold, so that confidence grows while capability stalls. The response is not to withhold support but to design its gradual withdrawal.

Self-efficacy and self-regulation are tightly coupled, since efficacious learners are more likely to set goals, monitor progress and persist (Sun & Wang, 2020; Zimmerman, 2002). Taught to use AI as a self-regulatory aid, prompting it to surface questions they should ask, learners may strengthen both confidence and the capacities that make confidence well-founded.

4.6. Homogenisation

A growing body of evidence suggests that widespread reliance on generative AI may homogenise writing. A controlled study found that access to generative AI raised the creativity of individual stories while reducing the collective diversity of the corpus they formed, so that gains at the individual level coincided with narrowing at the collective level (Doshi & Hauser, 2024).

For education, homogenisation poses a subtle but significant threat, because much of what writing instruction seeks to develop lies in the particular: the unexpected example, the personal angle, the argument that resists the obvious. Bandura's account of observational learning explains the mechanism, since learners who repeatedly emulate conventional output internalise its norms as the standard of good writing (Bandura, 1986).

The risk is uneven. For those with little prior writing experience the conventionalising influence may be especially strong, since they have not yet developed the personal repertoire from which to depart from a norm.

Homogenisation also carries an equity dimension. Generative systems are trained predominantly on text reflecting dominant linguistic and cultural conventions, and their outputs privilege certain registers and forms of argument. This is not speculative. Scholarship on academic writing in global context documents that Anglophone conventions of linearity, explicit thesis and deductive argument are neither universal nor neutral, and that writers from other rhetorical traditions are routinely required to assimilate to them (Canagarajah, 2002; Lillis & Curry, 2010). Research on second-language writers indicates that generative systems normalise text towards precisely those conventions, offering access to valued registers while eroding features that mark a writer's provenance (Warschauer et al., 2023). For multilingual writers, the pressure to conform may accelerate a loss of distinctive voice that writing education seeks to protect. The English-language limitation noted in Section 3.2 applies here with particular force.

Countering homogenisation requires positive effort rather than mere caution. Practices that compare AI-generated drafts with human alternatives, that ask learners to amplify what is distinctive in their own writing, and that explicitly value risk can preserve the diversity the technology tends to erode.

5. Discussion

Taken as a whole, the six themes describe a single underlying process: generative AI is altering writing by altering the cognition that produces it, within a reciprocal system that social cognitive theory illuminates. The environment, now populated by capable generative systems, changes what writing behaviours are easy and rewarded; those altered behaviours change the person, reshaping beliefs about capability and habits of monitoring; and the changed person re-enters the environment with new expectations. The influence runs in every direction, so neither technological determinism nor a faith in unaffected human autonomy describes the situation accurately.

The central argument can now be stated precisely. Generative AI changes writing by changing what writers believe they can do, what they choose to do, and what the environment makes easy. It changes belief by offering fluent production that raises confidence, potentially beyond the level capability warrants (Bandura, 1997; Huang & Mizumoto, 2024). It changes choice by making delegation the path of least resistance (Gerlich, 2025; Risko & Gilbert, 2016). And it changes the environment by establishing conventional AI-shaped prose as a default (Doshi & Hauser, 2024). These three limbs are not equally well evidenced. The claim about the environment rests on a controlled experiment. The claim about choice rests on cross-sectional and self-report evidence and is an association rather than a demonstrated mechanism. The claim about belief is a theoretically derived prediction that no study in the corpus has tested. In each case the effect is mediated by agency, and education is the practice through which agency can be defended or forfeited.

Generative AI is therefore best understood as neither a neutral tool nor an autonomous replacement for the writer, but as a mediating agent, in the restricted sense established in Section Agency, Proxy Agency, Technological Mediation and Affordance: it mediates the relation between intention and text and affords particular courses of action, without holding intentions or bearing responsibility. Where a learner secures a written outcome through the system rather than their own performance, the accurate description is proxy agency, in which the person retains the intention and the accountability while the capability is borrowed. This resists two errors: treating AI as inert, which the evidence on offloading

and homogenisation tells against, and treating it as a quasi-author, which the evidence on ownership tells against (Draxler et al., 2024; H.-P. Lee et al., 2025).

The decisive educational variable is critical awareness exercised through self-regulation. Agentic AI use is not the refusal of assistance but its deliberate, monitored and accountable employment, and the pattern across the corpus is consistent with this: studies associated with positive outcomes are those in which learners interrogate, verify and selectively incorporate output, while those associated with diminished thinking are those in which trust displaces critical engagement (Gerlich, 2025; H.-P. Lee et al., 2025), though both bodies of evidence are cross-sectional and self-reported. Two considerations carry this towards ethics. Evaluative judgement becomes more rather than less important when production is easy (Tai et al., 2018), since where fluent text is cheap the scarce capacity is the ability to judge whether it is true, apt and fit for purpose. And responsibility, which social cognitive theory frames as self-reactive influence and moral agency (Bandura, 2018), risks being diffused along with authorship unless deliberately reclaimed.

This is the point at which the hybrid position introduced in Section 1 can be restated in mature form, since the conclusion this review reaches follows from it. If human and machine contributions are genuinely entwined and difficult to separate (Draxler et al., 2024; Markauskaite et al., 2022), the significant question is not how to disentangle them but who is directing the entwinement and to what end. Hybridity does not distribute responsibility across the human and the system, because only one party can hold an intention or answer for a claim. The consequence is the opposite of the one sometimes drawn from hybrid accounts: the more entwined composing becomes, the more the deliberate and accountable participation of the human writer must be cultivated, precisely because it is no longer guaranteed by the mere fact of having produced the words.

A distinction long drawn in research on cognition and technology sharpens the stakes. Salomon and colleagues differentiated the effects with a technology, the performance achieved while using it, from the effects of a technology, the residue remaining after it is set aside (Salomon et al., 1991). Much of the evidence documents effects with the tool. The decisive question concerns the effects of the tool, and here the evidence is thinner and more equivocal than the confidence of current debate suggests: the studies that speak to it are short and cross-sectional, and they raise rather than settle the possibility that reliance leaves little durable residue (Gerlich, 2025; Kosmyna et al., 2025). Education is being asked to act on a question the evidence cannot yet answer, and the study most needed, a longitudinal investigation tracking unaided capability alongside sustained AI use, does not exist. The aim is not to maximise performance with AI but to ensure its lasting effects are developmental rather than degrading.

The reciprocal system operates differently across sectors, and a single policy cannot fit them all. The governing principle is constant, that AI use should be tilted towards agency, but the means must suit the developmental situation of the learners, as Section 6.4 sets out.

6. Application to Education

6.1. Principles for Agentic Human-AI Writing

The analysis yields guidelines for educators across sectors, oriented towards preserving human agency in the passage from thinking to writing. The first principle is to teach AI as part of the writing process rather than a shortcut around it, so that generative systems are encountered as participants whose contributions are directed and judged. The third is to design tasks that preserve planning, judgement and revision as human responsibilities. The fourth is to build confidence without encouraging dependence, sequencing assistance so that learners accumulate the mastery experiences on which durable self-efficacy depends (Bandura, 1997). The sixth is to discuss bias, voice and authorship openly. The

seventh and overarching principle is to cultivate critical AI literacy (Long & Magerko, 2020; Ng et al., 2021).

These are proposals derived from the synthesis rather than practices validated by trial. Table 6 makes their derivation auditable by tracing each of the five pedagogical levers in Figure 1 to the themes that generated it and the sources supporting it. Several levers rest substantially on conceptual rather than empirical warrant, and no lever has been evaluated as an intervention, which is why Figure 1 is offered as an evidence-informed conceptual proposition rather than a validated framework.

Table 6. Derivation of the pedagogical levers from the thematic findings.

Pedagogical Lever	Themes from Which It Derives	Supporting Sources in the Corpus	Basis of Warrant
Build critical AI literacy	Metacognition; homogenisation; authorship	Long and Magerko (2020); Ng et al. (2021); Markauskaite et al. (2022)	Conceptual; some cross-sectional support
Preserve planning and revision	Cognitive offloading; writing processes	Risko and Gilbert (2016); Flower and Hayes (1981); M. Lee et al. (2022); Gerlich (2025)	Theoretical; associational evidence
Require disclosure and reflection	Metacognition; authorship	Draxler et al. (2024); Zimmerman (2002); Cotton et al. (2024)	Conceptual; extrapolated from ownership evidence
Develop evaluative judgement	Authorship; metacognition; writing processes	Tai et al. (2018); Shi and Aryadoust (2024); Su et al. (2023)	Empirical for feedback; conceptual for judgement
Foster confidence, not dependence	Self-efficacy; cognitive offloading	Bandura (1997); Huang and Mizumoto (2024); Song and Song (2023); Sun and Wang (2020)	Empirical for short-term efficacy; theoretical for dependence

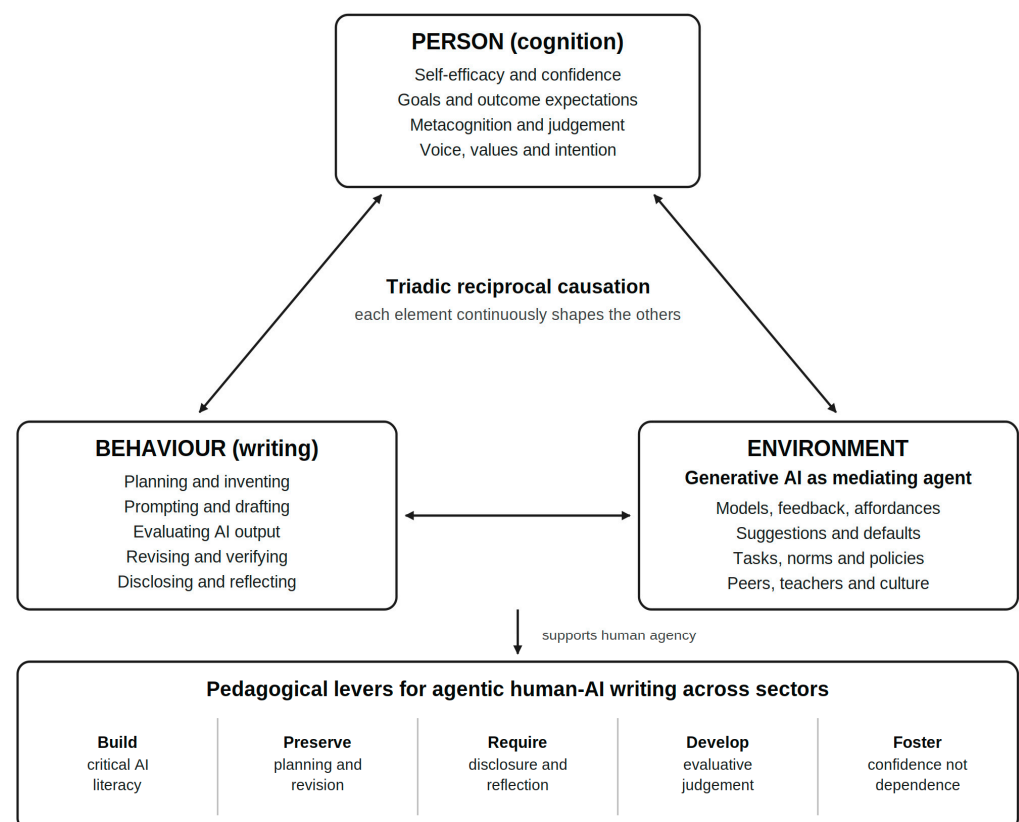


Figure 1. A social-cognitive model of agentic human-AI writing in education.

Note. The upper triad reproduces Bandura's triadic reciprocal causation, with the six themes distributed across its elements: self-efficacy and metacognition within the person;

writing processes, offloading and disclosure within behaviour; and generative AI, with the homogenising norms it carries, within the environment. The lower box was derived from the themes rather than appended to them: each theme was examined for the point at which a teacher could intervene, candidate levers addressing a common mechanism were grouped, and five were retained. Table 6 records this derivation lever by lever. Neither the model nor any lever has been empirically validated.

The second principle is to require disclosure and reflection. Documenting how AI was used, which suggestions were accepted or rejected, and why, makes a learner's agency visible and converts AI use into an occasion for metacognition rather than an unexamined transaction.

The fifth principle is to compare human and AI drafts explicitly, inviting learners to notice differences in voice, risk and distinctiveness and thereby to resist the homogenising pull of conventional machine prose (Doshi & Hauser, 2024).

6.2. A Worked Example: Disclosure and Reflection as a Classroom Task

The principles risk reading as aspirations unless one is shown in operational form. The following instantiates disclosure and reflection for a senior secondary or undergraduate assignment, and scales down for adult and vocational settings. Learners are told at the outset that AI may be used and that the assessed artefact has three parts. The first is the essay. The second is a process log of no more than 400 words, structured by four prompts: what I asked the system and why; what it gave me; what I kept, changed or rejected, and on what grounds; and what I would have done differently unaided. The third is a five-minute paired conversation in class, in which each learner explains one substantive change to machine output and defends it to a partner required to challenge the reasoning. Marking weights the essay at fifty per cent, the reasoning in the log at thirty, and the defence at twenty, and the rubric rewards identifying at least one instance where the learner judged the system wrong or ill-suited to the task.

Two features do the work. Grading the reasoning rather than the disclosure removes the incentive to declare minimal use. And pairing the log with a live defence means the log must be underwritten by knowledge the learner holds.

6.3. Making the Guidelines Verifiable

A serious objection applies here, and with particular force to disclosure and reflection. Any artefact a learner can be asked to produce, whether a process log, a reflective commentary or a planning document, can itself be generated by the technology it is meant to render accountable, and the result may be more plausibly self-aware than an honest account. Detection tools cannot resolve this, since their reliability is contested and their false positives fall unevenly on second-language writers (Cotton et al., 2024; van Dis et al., 2023). Guidelines resting on documentation alone are therefore vulnerable to the practice they exist to govern. The objection is well made and the original submission did not answer it.

The answer is not more elaborate documentation, and it is not a guarantee of fair use, which no assessment design has ever offered and which generative AI has made no less unattainable than it always was. What is available is a shift in what carries evidential weight. Documentation should be treated as a prompt for thinking rather than proof of it, and assurance should rest on conditions a learner cannot delegate. Four are available, none requiring surveillance. Real-time performance means short, supervised writing episodes without generative systems, establishing an unaided baseline and serving the calibration purpose identified in Section 4.2. Dialogic accounting means brief oral exchanges in which learners defend decisions in a text bearing their name, since machine-generated reflection does not survive an unrehearsed follow-up question. Visible process means

work developing across sessions with intermediate stages seen by the teacher. Contextual anchoring means tasks tied to material the system cannot access, such as a class discussion held that week, a placement, or the learner's own recorded observations.

None of these guarantees fair use, and it would be misleading to suggest otherwise. What they do is relocate assurance from what learners assert about their process to what they can demonstrate under observed, dialogic or contextually anchored conditions. The principal cost is teacher time, and that is the real constraint on adoption, most acutely in large cohorts and in the under-resourced adult and vocational sector where the enabling case for AI is strongest. Sampling offers a partial answer, since a defence with a rotating subset of a cohort changes the expectations under which everyone works. There is also a prior question of incentives: where an assessment can be satisfied by unaccountable delegation, learners will find the shortest route to it, and the more durable remedy is to set tasks whose completion requires the reasoning the course exists to develop. Establishing which combinations preserve agency at acceptable cost is a programme for assessment research this review can identify but not resolve.

6.4. *Adapting the Principles Across Sectors*

These principles require adaptation rather than uniform application. In schools, assistance is introduced sparingly so that core developmental work is left to the learner. In universities, the emphasis shifts to integrating AI into disciplinary writing while redesigning assessment to reward defensible judgement (Tai et al., 2018). In adult and vocational education, the priority is to honour the enabling potential of AI while sequencing its withdrawal so that autonomy develops; here the time cost of verification is most acute, and sampled dialogic checks are the most realistic instrument. What unifies these adaptations is critical AI literacy (Long & Magerko, 2020; Markauskaite et al., 2022; Ng et al., 2021).

7. Conclusions

Generative artificial intelligence is reshaping cognition and writing through new patterns of collaboration, delegation and dependence, most consequentially in education. This review has argued that the technology is best understood not as a neutral tool or an autonomous replacement but as a mediating agent within a social-cognitive system, a description denoting functional mediation and affordance rather than any attribution of intention. Read through Bandura's theory, the six themes describe the redistribution of cognitive and compositional labour between person and machine, with effects that depend on whether that redistribution is agentic or passive.

The evidence counsels neither uncritical enthusiasm nor reflexive prohibition, and it is more provisional than the urgency of current debate allows. Generative AI demonstrably supports confidence, access and efficiency, particularly for novice and multilingual writers. Whether the same affordances erode the capacities they assist is not settled: the evidence for erosion is associational, short-term and, in the most striking instance, not peer reviewed, and the divergence between confidence and competence remains a prediction rather than a finding. What can be said is that the difference between augmentation and erosion is a property not of the technology but of the human relationship to it.

The practical upshot follows from the hybrid account developed in Section 5. If human and machine contributions are entwined, responsibility does not divide with them, because only the human party can hold an intention or answer for a claim. The educational value of generative AI therefore depends less on access to tools than on the development of critical, agentic and reflective users, whose decisive capabilities are self-regulation, evaluative judgement and the disposition to engage AI deliberately. These are teachable,

and this article has set out both the principles through which they might be taught and the conditions under which such teaching can be verified.

The wider significance is that the question generative AI poses to education is not primarily technological but human. It asks what kinds of thinkers and writers we wish to cultivate, and whether the convenience of delegation will displace the formative work through which thinking and writing develop together. To think with the machine, rather than let the machine think in one's stead, is an achievement of agency that must be taught, modelled and practised.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijcs2030019/s1>, Table S1, verbatim search strings by information source, with fields, filters, dates and per-source yields; Table S2, characteristics and appraisal of the 36 included publications. The screening spreadsheet, extraction forms and analytic memos are available from the author on request.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The author declares no conflict of interest.

Statement about the Use of Generative Artificial Intelligence: Anthropic (Claude Opus 4.7) was used in the preparation of this article for the following purposes: assisting in editorial tasks, reducing word length, and checking compliance of the references. The same system was used during revision to check the internal consistency of the reported counts across the text, the tables and the supplementary evidence table. All searching, screening, appraisal, coding, interpretation and writing were performed by the author.

References

- Adams, R. J., Smart, P., & Huff, A. S. (2017). Shades of grey: Guidelines for working with the grey literature in systematic reviews for management and organizational studies. *International Journal of Management Reviews*, 19(4), 432–454. [\[CrossRef\]](#)
- Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: Methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-Based Healthcare*, 13(3), 132–140. [\[CrossRef\]](#) [\[PubMed\]](#)
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. [\[CrossRef\]](#) [\[PubMed\]](#)
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52(1), 1–26. [\[CrossRef\]](#) [\[PubMed\]](#)
- Bandura, A. (2018). Toward a psychology of human agency: Pathways and reflections. *Perspectives on Psychological Science*, 13(2), 130–136. [\[CrossRef\]](#) [\[PubMed\]](#)
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. [\[CrossRef\]](#)
- Canagarajah, A. S. (2002). *A geopolitics of academic writing*. University of Pittsburgh Press.
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19. [\[CrossRef\]](#)
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. [\[CrossRef\]](#)
- Critical Appraisal Skills Programme. (2018). *CASP qualitative studies checklist*. CASP UK.
- Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28), eadn5290. [\[CrossRef\]](#) [\[PubMed\]](#)
- Draxler, F., Werner, A., Lehmann, F., Hoppe, M., Schmidt, A., Buschek, D., & Welsch, R. (2024). The AI ghostwriter effect: When users do not perceive ownership of AI-generated text but self-declare as authors. *ACM Transactions on Computer-Human Interaction*, 31(2), 25. [\[CrossRef\]](#)
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911. [\[CrossRef\]](#)
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387. [\[CrossRef\]](#)

- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6. [\[CrossRef\]](#)
- Han, J., & Reinhardt, J. (2022). Autonomy in the digital wilds: Agency, competence, and self-efficacy in the development of L2 digital identities. *TESOL Quarterly*, 56(3), 985–1015. [\[CrossRef\]](#)
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388. [\[CrossRef\]](#)
- Hong, Q. N., Pluye, P., Fabregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O’Cathain, A., Rousseau, M.-C., & Vedel, I. (2018). *Mixed methods appraisal tool (MMAT), version 2018*. Canadian Intellectual Property Office.
- Huang, J., & Mizumoto, A. (2024). Examining the effect of generative AI on students’ motivation and writing self-efficacy. *Digital Applied Linguistics*, 1, 102324. [\[CrossRef\]](#)
- Hutchins, E. (1995). *Cognition in the wild*. MIT Press.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., . . . Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. [\[CrossRef\]](#)
- Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. *PsyArXiv*. [\[CrossRef\]](#)
- Latour, B. (1994). On technical mediation: Philosophy, sociology, genealogy. *Common Knowledge*, 3(2), 29–64.
- Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI conference on human factors in computing systems, Yokohama, Japan, 26 April–1 May 2025* (Article 1121). Association for Computing Machinery. [\[CrossRef\]](#)
- Lee, M., Liang, P., & Yang, Q. (2022). CoAuthor: Designing a human-AI collaborative writing dataset for exploring language model capabilities. In *Proceedings of the 2022 CHI conference on human factors in computing systems, New Orleans, LA, USA, 29 April–5 May 2022* (Article 388). Association for Computing Machinery. [\[CrossRef\]](#)
- Lillis, T., & Curry, M. J. (2010). *Academic writing in a global context: The politics and practices of publishing in English*. Routledge.
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–16). Association for Computing Machinery. [\[CrossRef\]](#)
- MacInnis, D. J. (2011). A framework for conceptual contributions in marketing. *Journal of Marketing*, 75(4), 136–154. [\[CrossRef\]](#)
- Markauskaite, L., Marrone, R., Poquet, O., Knight, S., Martinez-Maldonado, R., Howard, S., Tondeur, J., De Laat, M., Buckingham Shum, S., Gašević, D., & Siemens, G. (2022). Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI? *Computers and Education: Artificial Intelligence*, 3, 100056. [\[CrossRef\]](#)
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. [\[CrossRef\]](#)
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16(1), 1–13. [\[CrossRef\]](#)
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. [\[CrossRef\]](#) [\[PubMed\]](#)
- Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., & Koffel, J. B. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews*, 10(1), 39. [\[CrossRef\]](#) [\[PubMed\]](#)
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. [\[CrossRef\]](#) [\[PubMed\]](#)
- Salomon, G. (Ed.). (1993). *Distributed cognitions: Psychological and educational considerations*. Cambridge University Press.
- Salomon, G., Perkins, D. N., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher*, 20(3), 2–9. [\[CrossRef\]](#)
- Shi, H., & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. *ReCALL*, 36(2), 187–209. [\[CrossRef\]](#)
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. [\[CrossRef\]](#)
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1260843. [\[CrossRef\]](#) [\[PubMed\]](#)
- Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, 333(6043), 776–778. [\[CrossRef\]](#) [\[PubMed\]](#)
- Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57, 100752. [\[CrossRef\]](#)

- Sun, T., & Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System*, 90, 102221. [[CrossRef](#)]
- Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, 76(3), 467–481. [[CrossRef](#)]
- Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, 4(3), 356–367. [[CrossRef](#)]
- van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, 614(7947), 224–226. [[CrossRef](#)] [[PubMed](#)]
- Verbeek, P.-P. (2005). *What things do: Philosophical reflections on technology, agency, and design*. Pennsylvania State University Press.
- Wang, C. (2024). Exploring students' generative AI-assisted writing processes: Perceptions and experiences from native and nonnative English speakers. *Technology, Knowledge and Learning*, 30, 1825–1846. [[CrossRef](#)]
- Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing*, 62, 101071. [[CrossRef](#)]
- Wertsch, J. V. (1998). *Mind as action*. Oxford University Press.
- Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, 52(5), 546–553. [[CrossRef](#)] [[PubMed](#)]
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.