Journal of Transformative Pedagogies and Learner Engagement

Vol. 1. **Issue** 3, 2026. Pp. 1-11
ISSN: 3139-2784

Journal homepage: https://spublication.in/journals/index.php/jtple



Article

Borrowed Authorship: AI-Mediated Academic Writing, Authorial Disconnect, and the Problem of Intellectual Ownership

Ghulam Mohammad Khan[1*], Manzoor Ahmad Parey[1], Mubashir Hamid Mir[1], Irfan Ahmad Wagay[1]

1. HKM Govt. Degree College Bandipora, Higher Education Department, J&K, India.

*Corresponding Author:
gluoomkk@gmail.com

Article History

Received:	26/03/2026
Revised:	20/08/2026
Accepted:	06/09/2026
Published:	30/09/2026

DOI: 10.63960/sp.2026.jtple.1.3.32

Abstract:

The rapid integration of generative artificial intelligence (AI) into academic writing has directed scholarly attention to a single question: was a given text produced with the help of a machine? This paper argues that a more consequential question concerns the author rather than the artefact: does the person whose name appears on a published paper understand, and remember, what the paper says? The study began as a hypothesis formed during a routine discussion among three faculty members at a degree college in North Kashmir. To test it, we identified twenty already-published research papers across the sciences, social sciences, and humanities: ten flagged by established AI-detection tools as having more than 80 per cent AI usage (an imperfect proxy for actual AI use), and ten showing no AI assistance. We then conducted voluntary, semi-structured interviews with first authors, preferably by telephone so that authors could not immediately revisit their papers, and compared their command of each paper's concepts, terminology, borrowed ideas, methodology, and contribution to the field. Nine of ten authors of AI-mediated papers showed a marked disconnect from their own content, despite the fluency, coherence, and argumentative strength of the texts. Eight of ten authors who wrote without AI answered readily and confidently and clearly recalled the experience of writing and publishing. The AI-mediated papers also tended to appear in less popular journals, and five of ten contained mismatches in the references cited. These are exploratory associations from a small purposive sample and do not establish that AI use caused the differences. Interpreting these patterns through theories of authorship, levels of processing, cognitive offloading, and processing fluency, we propose the concept of borrowed authorship: a condition in which ownership of a text and possession of the knowledge it contains come apart. We conclude that research integrity regimes should move from detecting machine-generated text towards assessing authorial comprehension.

Keywords: AI-Mediated Writing, Authorship, Authorial Disconnect, Cognitive Offloading, Memory, Research Integrity, Metacognition, Academic Publishing

1. INTRODUCTION

Every technology of writing has provoked an anxiety about knowing. In the *Phaedrus*, Socrates recounts the myth of Theuth, the inventor of writing, who presents his discovery to King Thamus as a remedy for memory and wisdom. The king is unmoved. Writing, he replies, will produce forgetfulness in the souls of those who learn it, because they will trust external marks instead of remembering from within, and they will acquire the appearance of wisdom rather than its reality (Plato, 1995). Derrida (1981) observed that the Greek word for Theuth's gift, *pharmakon*, means both remedy and poison, and that the text cannot finally decide between the two. That undecidability is instructive for the present moment. Generative AI, which can summarise a literature, propose a framework, draft an

Journal of Transformative Pedagogies and Learner Engagement is a peer-reviewed open-access journal. © 2026 The Author(s). This is an open-access article published by Synergy Publication under the Creative Commons Attribution 4.0 International License (CC BY 4.0). This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and source are credited. The Author(s) retain copyright of their work. Synergy Publication is granted a non-exclusive right to publish the article and to identify itself as the original publisher. For more information, *See http://creativecommons.org/licenses/by/4.0/.*

OPEN ACCESS

argument, and polish a paragraph in seconds, is the most powerful pharmakon of writing yet devised. It is also, for many researchers, a quiet invitation to possess the products of understanding without undergoing the labour of it.

The academic response to this invitation has been largely regulatory. Leading journals and publishers have declared that AI systems cannot be listed as authors because they cannot take responsibility for a work (Thorp, 2023; Stokel-Walker, 2023), and editorial bodies now require authors to disclose AI assistance and remain accountable for the whole of a submission (International Committee of Medical Journal Editors [ICMJE], 2024). Universities and publishers have adopted detection software, even though independent testing has found such tools neither accurate nor reliable enough to bear the weight placed upon them (Weber-Wulff et al., 2023), and they have been shown to misclassify the writing of non-native English speakers (Liang et al., 2023). The debate has therefore been organised around the artefact: is this text machine-made, and can we prove it?

This paper starts from a different place. Whatever its provenance, a published paper is attributed to a person who owns the text and gains the benefits of authorship: recognition, citation, promotion, and professional standing. The modern concept of the author is, as Foucault (1977) argued, less a description of a writing individual than a function that organises discourse, assigns responsibility, and regulates ownership. Barthes (1977) had already proclaimed the author's death in order to liberate the reader. Yet the economy of scholarly credit has continued to depend upon a living, answerable author. The question that generative AI places before this economy is therefore not only whether the author wrote the text but whether the author *knows* it. What remains to be explored, in the words of the hypothesis that prompted this study, is how much the author remembers clearly from the AI-mediated text.

The hypothesis arose in a routine conversation among the three of us, faculty members at a degree college in North Kashmir, about the integration of AI into academia and the changing position of the author. We suspected that, owing to overreliance on AI and distance from the textual reading material, a noticeable disconnect had developed between the text, or the research, and its author. We emphasise that this suspected disconnect does not depend on the text being detected as AI-mediated, assisted, or generated. It depends on the author being largely disconnected from, or even unaware of, the content of the research that carries their name. To explore whether the hypothesis could be found true, we examined published papers and spoke to their first authors.

The study was guided by three research questions:

RQ1. To what extent do authors of AI-mediated papers understand and recall the concepts, terminology, borrowed ideas, methodology, and contribution of their own published work, compared with authors who wrote without AI assistance?

RQ2. How do authors in the two groups narrate the experience of writing and publishing their papers?

RQ3. Do the two groups differ, as a preliminary observation, in the publication venues they reach and in the accuracy of the references they cite?

The question has particular salience in contexts where publication is tied to promotion and institutional recognition, and where scholars working at the periphery of global knowledge economies face pressure to produce internationally visible work with limited institutional support (Connell, 2007). In such contexts AI tools may appear as an equaliser, and in some respects they are. The findings reported here suggest, however, that the equalisation of *text* may conceal a widening inequality of *understanding*. The remainder of the paper reviews the relevant literature and sets out a conceptual framework, describes the method, presents the findings, discusses their significance, acknowledges the study's limitations, and concludes with implications for authors, editors, institutions, and educators.

2. LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK

2.1. Authorship: ownership, function, and accountability

Theories of authorship have long separated the individual who writes from the authority who is credited. Barthes (1977) argued that a text is a tissue of quotations drawn from innumerable centres of culture, and that the unity of a text lies in its destination, the reader, rather than its origin. Foucault (1977) accepted the erosion of the romantic author but insisted on the persistence of the *author-function*, a mechanism by which discourses are classified, ownership is assigned, and punishment or praise is made possible. Scholarly publishing is the clearest institutional embodiment of this function. The author's name stabilises the text, links it to a career, and makes someone answerable for its claims.

Contemporary editorial standards make that answerability explicit. The ICMJE (2024) criteria require not only a substantial contribution to the conception, design, or analysis of the work and its drafting or critical revision, but also agreement to be accountable for all aspects of the work, ensuring that questions about its accuracy or integrity are appropriately investigated and resolved. Accountability of this kind presupposes comprehension, since a person who cannot explain a paper's concepts or methods cannot answer for them. The authorship norms that exclude AI systems on the ground that they cannot be accountable (Thorp, 2023; Stokel-Walker, 2023) thus leave an assumption in place: that the human author, by contrast, can be. The present study tests this assumption empirically.

2.2. Memory, externalisation, and cognitive offloading

The anxiety expressed in the *Phaedrus* has a long afterlife. Ong (1982) traced how the shift from orality to literacy reorganised thought, making possible abstraction and analysis while weakening the mnemonic disciplines of oral cultures. Stiegler (1998) described technical supports for

https://doi.org/10.63960/sp.2026.jtple.1.3.32

memory, which he called *hypomnemata*, as constitutive of human cognition but also as pharmacological: they extend memory and at the same time threaten to dispossess individuals of the knowledge they externalise. Carr (2010) popularised a related concern about the internet's effect on sustained attention and deep reading.

Experimental psychology has given these concerns empirical form. Sparrow, Liu, and Wegner (2011) showed that when people expect information to remain accessible through a computer, they tend to remember where to find it rather than the information itself. Wegner's (1987) account of transactive memory, in which groups distribute memory across members, helps explain why this can be adaptive. The key condition is that the partner holds the memory reliably and that the individual knows what the partner holds. Risko and Gilbert (2016) reviewed the broader phenomenon of cognitive offloading, concluding that offloading can improve performance in the moment while altering what is encoded and retained. The generative AI case differs from the search-engine case in an important respect. What is offloaded is not merely facts but interpretive labour, namely the work of reading a source, deciding what it means, relating it to other sources, and articulating the result. If memory follows engagement, as the research discussed below suggests, the offloading of interpretation should affect what the author later remembers about the research.

2.3. Depth of engagement, generation, and tacit knowledge

The levels-of-processing framework holds that durable memory depends on the depth at which material is processed: semantic, effortful processing yields more lasting traces than shallow, perceptual processing (Craik & Lockhart, 1972). The generation effect shows that information a person produces for themselves is remembered better than information they merely read (Slamecka & Graf, 1978). Bjork and Bjork (2011) argued that conditions which make learning harder in the short term, which they call desirable difficulties, often enhance long-term retention and transfer. Chi and Wylie (2014) distinguished passive, active, constructive, and interactive modes of engagement and showed that learning outcomes rise as engagement moves from receiving to generating and co-creating knowledge. Delegating the drafting of an argument to an AI system moves the researcher towards the passive or at best the evaluative end of this continuum.

The humanities and the philosophy of knowledge point in the same direction. Rosenblatt (1978) described reading as a transaction in which the reader and the text jointly produce the poem, so that meaning is an event rather than an object. Ricoeur (1981) spoke of *appropriation*, the process by which a reader makes foreign meaning one's own and is transformed in the process. Polanyi (1966) argued that we know more than we can tell, and that this tacit dimension of knowledge is acquired by dwelling within a body of material until it becomes an extension of one's own perception. A researcher who has read widely, struggled with a method, and drafted and redrafted an argument carries tacit knowledge that no summary can substitute. A researcher who receives the finished synthesis acquires the proposition but not the dwelling.

2.4. Fluency, confidence, and the monitoring of one's own understanding

Flavell (1979) defined metacognition as knowledge of and control over one's own cognitive processes, and a central part of it is monitoring: judging accurately what one does and does not know. Monitoring is vulnerable to processing fluency, the subjective ease with which information is processed. Alter and Oppenheimer (2009) reviewed evidence that fluency shapes judgements of truth, familiarity, and confidence, and Kahneman (2011) located the effect in the quick, associative operations of System. Rozenblit and Keil (2002) demonstrated an *illusion of explanatory depth*: people believe they understand how things work far better than they do until asked to explain. Generative AI produces maximally fluent text. Reading such text may therefore generate a feeling of understanding that is not tested until the author is asked to explain, defend, or apply what has been written.

Research on human interaction with automation adds a behavioural dimension. Parasuraman and Manzey (2010) documented complacency and automation bias, in which operators reduce their monitoring of automated systems and accept their output uncritically. A systematic review of automation bias in clinical decision support reached similar conclusions (Goddard, Roudsari, & Wyatt, 2012). For writers using AI, complacency may take the form of accepting a fluent draft, its claims, and its citations without verification.

2.5. The unreliability of generative output and of its detection

Large language models generate statistically plausible text and are known to produce hallucinations, content that is fluent but unfaithful to sources or fabricated (Ji et al., 2023). Bender et al. (2021) cautioned that such systems manipulate linguistic form without access to meaning, and Hicks, Humphries, and Slater (2024) argued that their outputs are better described as indifferent to truth than as mistaken. The problem has a direct bibliographic form. Walters and Wilder (2023) found that a substantial share of the citations generated by ChatGPT were fabricated and that many of the remainder contained errors. A researcher who has not read the cited sources cannot easily detect such errors, which makes the accuracy of a reference list a useful trace of the author's contact with the literature.

Detection tools, meanwhile, remain unreliable. Weber-Wulff et al. (2023) tested a range of detectors and found them inaccurate and easily degraded by simple paraphrasing, and Liang et al. (2023) found that they disproportionately flagged the writing of non-native English speakers as machine-generated. Any study that relies on detectors to classify texts must therefore treat their output as a pragmatic, imperfect indicator and must triangulate it. We return to this point in the methodology and limitations.

2.6. Emerging evidence on AI use and cognition

A growing body of work examines how generative AI affects cognition. Kasneci et al. (2023) surveyed the opportunities and challenges of large language models in education, including the risk of overreliance and diminished critical thinking. Gerlich (2025) reported that more frequent AI tool use was associated with lower critical-thinking scores, mediated by cognitive offloading. Lee et al. (2025), surveying knowledge workers, found that higher confidence in AI was associated with less critical thinking, while higher self-confidence was associated with more. Kosmyna et al. (2025), in a preprint using neurophysiological measures, reported weaker neural engagement and poorer recall of their own essays among participants who wrote with a large language model. These studies rely largely on self-report, short-term tasks, or student samples, and some are not yet peer reviewed. What is missing is evidence about *published researchers* and their recall of *their own published work*. The present study addresses that gap in an exploratory way.

Recent research has also examined the broader attitudes, practices, and ethical implications associated with the integration of AI into academic research. Asagar (2026), in a quantitative descriptive study of postgraduate students and research scholars, investigated awareness, patterns of AI utilisation, perceived benefits, and ethical concerns surrounding AI-assisted research. The findings indicated that although respondents demonstrated considerable familiarity with AI tools and recognised their potential contribution to research quality, important gaps remained in practical utilisation, formal training, and ethical preparedness. Concerns relating to plagiarism, algorithmic bias, data privacy, intellectual ownership, and over-reliance on AI were also evident. While Asagar's study primarily examines perceptions and practices of AI adoption, the present investigation extends this discussion by exploring a more fundamental question: whether the use of AI in academic writing may create a separation between textual authorship and the author's actual comprehension, memory, and intellectual engagement with the published work.

2.7. Conceptual framework: from engagement to answerable authorship

Drawing on the literature above, we propose a simple chain through which authorship ordinarily operates. Sustained engagement with sources and data (reading, interpreting, drafting, revising) produces comprehension and durable memory of the content. Comprehension and memory in turn sustain confidence, the ability to explain, defend, and situate the work, and therefore accountable authorship. AI mediation inserts a machine between the researcher and the stages of engagement. The text that results may be excellent, but the intermediate cognitive work has been done elsewhere. We use the term *borrowed authorship* for the resulting condition, in which an author holds title to a text without holding the understanding that would ordinarily accompany it. In the language of one of our field observations, lived experience becomes borrowed experience. The study examines whether this condition is observable in practice, and what traces it leaves in recall, confidence, publication venue, and referencing.

3. METHODOLOGY

3.1. Research design

The study is an exploratory comparative qualitative inquiry with descriptive quantitative elements. Its logic is abductive. The hypothesis was generated from the researchers' professional experience and informal discussion, and the empirical work was designed to subject it to a first, disciplined test. A two-group comparison was chosen so that observed patterns could be read against the general difficulty of recalling one's own published work, since all authors, whatever their method, can forget. The unit of analysis was the paper and its first author. Participants were not randomly assigned, AI use was not independently verified, and potential confounding variables were not controlled. The study can therefore identify associations and generate propositions, but it cannot establish that AI mediation caused the differences observed.

3.2. Sampling and classification of papers

We used purposive sampling to construct two groups of ten already-published research papers each. The first group comprised papers with more than 80 per cent AI usage, and the second comprised papers from the same fields with no AI assistance. Papers were drawn from the sciences, social sciences, and humanities so that the findings would not reflect the conventions of a single discipline. The assignment of identifying and selecting papers was distributed among the three researchers. The two groups were not matched on discipline, publication year, author career stage, number of co-authors, or language background, and replacing authors who declined to participate may have introduced volunteer-selection bias.

To classify the papers we used AI-detection tools generally regarded as trusted: Originality.AI, Turnitin AI, QuillBot, Winston AI, Grammarly AI, and GPTZero. We recognise that such tools are imperfect proxies (Weber-Wulff et al., 2023). Using several of them, rather than one, was intended to reduce the risk that a single tool's idiosyncrasies would determine group membership. Because the study concerns the author's relationship to the text rather than the text's detectability, the classification served only to locate papers likely to differ in the degree of AI mediation. A detection score is not a direct measure of actual AI use. It differs from the proportion of text actually generated by AI and from the extent to which AI contributed to a paper's conceptual development, argument, or language editing, and a low score does not prove that AI was not used. The terms "AI-mediated" and "without AI" are therefore used in this paper to denote classification by detector output and not verified AI use. The central evidence comes from the interviews. The papers selected for the study were published during and after 2023.

3.3. Data collection

After selecting the papers, we conducted an interaction with each first author, preferably over the telephone. The medium was chosen deliberately. A telephone conversation does not allow authors to revisit their papers immediately, so the responses reflect what the author actually knows and

https://doi.org/10.63960/sp.2026.jtple.1.3.32

remembers rather than what can be recovered by rereading. The interactions were voluntary and not compulsory. Where an author was reluctant or declined, a new paper was identified in its place so that the sample reflected willing participants only.

Before each interaction we studied the paper carefully and identified its key concepts, the theoretical methods employed, its terminology, the names of the scientists, theorists, and thinkers whose ideas were used or approximated, and the methodologies adopted. From these we formulated questions for the author. The questions covered five areas: the central concepts and terminology; the ideas drawn from other authors; the methodology; the gap or need that the study addressed; and the author's experience of writing the paper and choosing where to publish. In some cases we deliberately said that the paper was relevant to our own study and therefore needed clarification. This framing made the conversation natural for the author and avoided signalling that their recall was being tested. Because the questions were paper-specific, their difficulty varied across authors, which is a limitation. We also recognise that failing to recall details without reference to a paper does not automatically indicate a lack of understanding, since unaided recall, conceptual comprehension, and the ability to defend one's research decisions are related but distinct.

3.4. Data analysis

Responses were documented and compared across the two groups. We assessed each author's response along several dimensions: the accuracy and completeness of their explanation of the paper's concepts and terminology; their ability to identify and place borrowed ideas and thinkers; their account of the methodology; their articulation of how the study filled a gap; their apparent confidence; and their recall of the experience of writing and publishing. Patterns across responses were identified through thematic comparison in the manner described by Braun and Clarke (2006), and the frequency of each pattern within each group was counted. In addition, we compared the journals in which the papers appeared and examined the references cited in the papers. The three researchers discussed their readings of each case until they reached a shared interpretation. Responses were judged against the areas of questioning described above and were not scored with a formal rubric. The researchers shared the initial hypothesis and knew each paper's group, which creates a risk of confirmation bias: hesitation, for example, could be read as disconnection in one author and as ordinary forgetting in another. The frequency counts are descriptive only and are not evidence of statistical significance.

3.5. Ethical considerations

Participation was voluntary, and no author was pressured to take part. Authors who declined were replaced. Findings are reported in aggregate, and no individual author, paper, or journal is identified here. Because the framing of the interaction did not foreground the study's specific hypothesis, we regard the protection of participants' identities and the avoidance of any evaluative judgement of individuals as especially important. The limited non-disclosure was intended to avoid prompting authors to consult their papers before the interaction, and responses were not communicated to authors' institutions. The purpose of the research is to understand a structural condition of contemporary scholarship, not to evaluate or shame particular researchers.

3.6. Trustworthiness

Following Lincoln and Guba (1985), we sought credibility through the use of several detection tools, the participation of three researchers in selection and interpretation, and a comparison group that provided a baseline against which the AI-mediated group could be read. We do not claim methodological triangulation, since the detectors measure the same contested construct and discussion among researchers is not independent verification. Dependability and transferability are supported by the description of the sampling and interview procedures, although the sample is purposive and small. The consistency of patterns across the sciences, social sciences, and humanities provides additional support, since a finding that recurs across fields is less likely to be an artefact of one discipline's norms. The sample is small, however, and the study should be read as hypothesis-sharpening rather than hypothesis-confirming.

4. FINDINGS

After careful evaluation of the papers and the interactions with their authors, a strange and consistent pattern was observed across the responses from both categories. Table 1 summarises the principal findings, which are elaborated in the subsections that follow.

Table 1. *Summary of principal findings (n = 10 papers per group)*

Dimension	AI-mediated papers	Papers without AI	Observation
Command of concepts, terminology, borrowed ideas, methodology, and contribution	Marked disconnect: 9 of 10 authors	Almost correct, confident responses: 8 of 10 authors	Language and argument strong in AI-mediated papers; authors uncertain
Recall of writing experience and publication process	Not a feature of responses	Clearly recalled and elaborated	Unassisted authors narrated how the paper came to be
Publication venue	Less popular journals	Comparatively popular journals	Judged by field visibility, not formal metrics

Dimension	AI-mediated papers	Papers without AI	Observation
Mismatches in references cited	5 of 10 papers	Not reported	Types of mismatch not systematically coded

4.1. Authors of AI-mediated papers: fluent texts, uncertain authors

Among authors who had used AI to assist their writing or to generate content, there was a striking similarity of responses. Nine of the ten respondents showed an alarming disconnect from the content of their own papers. The disconnect was not confined to a single area. It extended to the concepts on which the paper turned, the terminology it employed, the critical ideas taken from other authors, the methodology used, and how the study filled a gap in the literature. When asked to explain a central concept, the authors tended to hesitate or to answer in generalities that did not match the specificity of the published text. When asked about the thinkers or scientists whose ideas the paper had used or approximated, they were often unable to say who they were, what they had argued, or why their work had been drawn upon. When asked to describe the methodology, they could not reconstruct the steps or the rationale. Only one of the ten did not display this pattern. This exception matters: it suggests that a high detection score does not necessarily accompany weak authorial comprehension, and that how AI is used may matter more than whether it is used. The present data do not allow this case to be analysed further.

The significance of this finding lies in a paradox. The language of these papers, their coherence, and their arguments were very strong. By any conventional editorial or peer-review standard they were accomplished texts. Yet their authors were largely uncertain about their own explorations. The text was articulate, and the author was not. The authors could claim the content, since it appeared under their names, but did not seem to be fully aware of what the content was all about.

4.2. Authors of unassisted papers: recall, confidence, and the memory of practice

The pattern among authors who did not use AI was strikingly different. Eight of the ten readily produced almost correct responses to the questions and exhibited a confidence that was lacking among the authors who had used AI. They answered promptly and moved easily between the paper's concepts, its sources, and its methodology. Their recall was not merely propositional. These authors clearly remembered the experience of writing the papers. They described using concepts from different thinkers, explained the context of the study, and said how it was new. They also remembered the purpose behind publishing and the reasons for selecting particular journals. While talking about their papers, they shared their experience and knowledge of the publication process: the submission, the review, and the revisions that shaped the final text.

The difference was therefore not simply that one group was more accurate. The unassisted authors possessed a narrative of the research, a story of how the paper came to be, in which concepts, sources, decisions, and publication strategy were interlinked. The AI-mediated authors, by contrast, tended to lack this narrative. The observation suggests that the more an author is involved with the study of the texts during research, the more likely are conceptual clarity and clear remembrance of the same.

4.3. Publication venues

The publications also showed a pattern in venue. The AI-mediated papers were published in less popular journals, whereas the papers that did not use AI were published in comparatively popular journals. We use the term *popular* in the sense of visibility and standing within the relevant field rather than as a formal metric. The difference was observed across the fields studied, though we did not compute indicators such as impact factors. This is therefore a subjective and preliminary observation, since explicit criteria for journal standing were not applied.

4.4. Mismatches in references

Finally, five of the ten AI-assisted papers showed mismatches in the references cited in the paper. The mismatches were relatively few in each case, but their occurrence in half of the AI-assisted sample is notable, because reference lists are one place where an author's actual contact with the literature leaves a visible trace. We did not systematically code the types of mismatch (for example, fabricated references, incorrect bibliographic details, incorrect attribution, or references that do not support the claim made), and an equivalent check is not reported for the unassisted group. This asymmetry means the observation is preliminary.

4.5. Summary of the pattern

Taken together, the findings describe two contrasting relationships between author and text. In the unassisted group, authorship coincided with knowledge: the author could explain, recall, and narrate. In the AI-mediated group, authorship coincided with ownership but frequently not with knowledge: the text was strong, the author was uncertain, and the traces of the writing process, including the references, showed signs of weak contact with the sources.

5. DISCUSSION

5.1. The authorship paradox: ownership without comprehension

The central pattern observed is an apparent dissociation, in the AI-mediated group, between ownership of a text and demonstrated knowledge of its content. Editorial policy responds to AI by locating accountability in the human author (Thorp, 2023; ICMJE, 2024). But accountability requires comprehension, and our findings suggest that comprehension cannot be assumed. If nine of ten authors in an AI-mediated sample cannot readily explain their own concepts, sources, and methods, then the normative

https://doi.org/10.63960/sp.2026.jtple.1.3.32

image of the accountable author may, in these cases, be more assumed than demonstrated, sustained by the author-function (Foucault, 1977). The author's name continues to organise discourse, assign credit, and allocate responsibility, while the cognitive substance that would ordinarily justify that name has been displaced.

The observation also complicates the celebratory reading of Barthes (1977). In his account the death of the author frees the text from the tyranny of a single origin. In the AI-mediated condition, the text is indeed freed from the author, but not in the interest of the reader. The institutional economy of credit still requires an author, and so a function is retained without the subject. It is as if the author-function has been separated from the author's knowledge. We note, as the AI-mediated authors' responses implied, that AI can create many authors, each owning a text and none necessarily inhabiting it.

5.2. From lived to borrowed experience

The contrast between the two groups is intelligible in terms of the literature on engagement and memory. The unassisted authors had done the work of processing the material deeply (Craik & Lockhart, 1972), had generated their own formulations (Slamecka & Graf, 1978), and had worked through difficulties that the AI-mediated authors were spared (Bjork & Bjork, 2011). In Chi and Wylie's (2014) terms, they had operated in the constructive and interactive modes, whereas the AI-mediated authors had more often operated in the passive or evaluative mode. We would expect the first mode of engagement to be remembered better. The narratives that the unassisted authors offered, including the purposes of publication and the choices of journals, are the kind of integrated, episodic memory that arises from sustained involvement.

Beyond the cognitive account lies a hermeneutic one. For Ricoeur (1981), understanding is appropriation: the reader makes the text's meaning their own and is changed by it. For Polanyi (1966), knowing is a tacit, embodied achievement that comes from dwelling within the material. When AI reads, interprets, and synthesises on the researcher's behalf, the process of appropriation takes place in the machine and not in the person. The researcher receives the output of an interpretation but not the transformation that interpreting ordinarily brings. In this sense, the lived experience of research becomes a borrowed experience. The author relies on AI to unravel the text's meanings and, in the process, distances themselves from the text. The lesser the engagement, the shorter-lived is the author's memory of the text. This is a possible mechanism suggested by the literature and by our field observation; the interview evidence is consistent with it but cannot establish it.

5.3. Fluency, the illusion of authorial depth, and the pharmakon

Why might the AI-mediated authors have been uncertain rather than merely forgetful? The research on fluency offers one possible explanation. Text generated by large language models is highly fluent, and fluent text invites the feeling that one understands it (Alter & Oppenheimer, 2009; Kahneman, 2011). When an author reads a clear, well-organised draft on their own topic, the ease of reading may be mistaken for the ease of knowing, a version of the illusion of explanatory depth (Rozenblit & Keil, 2002). The illusion is exposed only when the author is asked to explain, as our telephone interviews required. In metacognitive terms (Flavell, 1979), the authors' monitoring of their own understanding had been calibrated against the quality of the text rather than the quality of their knowledge. We might tentatively call this a possible *illusion of authorial depth*, although the present design cannot test it.

The interpretation echoes the ancient warning in the *Phaedrus* that writing gives its users the appearance of wisdom without its reality (Plato, 1995). Derrida (1981) reminds us that this pharmakon is both remedy and poison, and our findings do not condemn the technology. Nothing in the data suggests that AI cannot assist a researcher who reads, interprets, and verifies. Stiegler's (1998) pharmacological perspective points to the real issue: whether externalised memory and labour are accompanied by practices of care that keep the individual a participant in knowledge rather than its recipient. The unassisted authors in our sample were participants. Many of the AI-mediated authors appear to have been recipients.

5.4. What is offloaded, and where the memory lives

Studies of cognitive offloading suggest that people remember where information is rather than the information itself when they expect it to remain available (Sparrow et al., 2011), and that such offloading can be rational (Risko & Gilbert, 2016). Transactive memory works when a partner reliably holds what the individual does not (Wegner, 1987). The AI-mediated authors in our study, however, were not in a transactive relationship with a stable repository. A conversation with a generative model does not preserve the reasoning in a form the author can later consult, and the model does not remember the research on the author's behalf. The memory of the paper thus lives nowhere in particular. It is neither in the author's mind nor in a durable external record that the author can navigate. This differs from earlier technologies of memory, where the author at least produced the inscription. It helps explain why the disconnect we observed took the form of uncertainty about the content itself and not merely uncertainty about where it could be found.

5.5. Venues and references as traces of reading

The venue pattern should be interpreted with caution. Several explanations are plausible and compatible with one another. Editors and reviewers in higher-visibility journals may be more alert to superficial argumentation, derivative framing, or weak engagement with the literature, even when the prose is polished. Authors who are not deeply engaged with their material may be less able to respond to critical review, and so may be steered towards venues with lighter review. Alternatively, researchers working with fewer resources, or under intense pressure to publish, may be more likely to use AI and also more likely to publish in less visible journals, in which case AI use is a marker

of structural conditions and not their cause (Connell, 2007). Our design cannot separate these possibilities. What it does show is that the papers whose authors were most disconnected from them were not found in the most visible venues, which may suggest, though our data cannot confirm it, that venue standing is associated with the factors discussed here.

The reference mismatches are more directly interpretable. Language models are prone to producing plausible but unfaithful content (Ji et al., 2023), including fabricated or erroneous citations (Walters & Wilder, 2023). An author who has not read the sources is poorly placed to catch such errors, and an author who trusts the output is subject to the complacency that Parasuraman and Manzey (2010) and Goddard et al. (2012) describe. References are, in this sense, a forensic index of reading. A reference list is a record of the sources an author has consulted, and mismatches in it signal that the chain between author and literature has been interrupted. That half of the AI-assisted papers displayed such mismatches is, as a preliminary observation, consistent with the interview pattern, although an equivalent check of the other group would be needed to draw any firm inference.

5.6. Alternative explanations and caution

The associations reported here admit explanations other than AI use. An author's ability to explain a paper may depend on the time since publication, the number of co-authors, career stage, actual contribution to the work, language comfort during the interview, and the complexity of the paper. Poor recall is not the same as poor understanding: an author may have contributed substantially yet struggle to recall details later, and an author may remember terminology without having developed the argument independently. A high detection score is not evidence that AI produced the intellectual content. Causation may also run the other way, since authors with weaker command of a field may be more likely to use AI, and authors under publication pressure or with fewer resources may be both more likely to use AI and more likely to publish in less visible journals. These variables were not controlled, so the differences between the groups should not be attributed exclusively to AI mediation.

5.7. Implications

For authors. The findings do not require abstention from AI, but they suggest a reading-first discipline. Authors who use AI should read the primary sources themselves, draft the argument in their own words before consulting a model, use AI to critique and extend rather than to originate, and verify every reference. A practical test is whether the author could present and defend the paper to a sceptical colleague without notes.

For editors and publishers. Detection-centred policies attend to the wrong object. Given the unreliability of detectors (Weber-Wulff et al., 2023; Liang et al., 2023), and the possibility of false accusations, journals might shift from asking whether a text is machine-generated to asking whether the author can answer for it. Possible measures include structured statements of each author's intellectual contribution, routine verification of a sample of references, and brief author responses to reviewer queries that demand more than a polished paragraph. Such measures also protect honest authors, including non-native English writers whose prose is more likely to be flagged.

For institutions. Promotion and recognition systems that reward publication counts create incentives for text production decoupled from understanding. Institutions might supplement counting with practices that evaluate comprehension, such as oral presentations of published work, research seminars, and doctoral vivas that probe the sources and methods of articles submitted in support of a thesis.

For teaching and supervision. If dependence on AI weakens recall and confidence, as this study and the wider evidence suggest (Gerlich, 2025; Lee et al., 2025; Kosmyna et al., 2025), then pedagogy should reinstate productive struggle. Students and early-career researchers need structured experiences of reading closely, summarising from memory, and explaining their work orally, because these are the practices through which tacit knowledge is built (Polanyi, 1966; Bjork & Bjork, 2011).

For future research. Larger and more systematically sampled studies are needed, including comparisons of recall by the extent of AI use, controlled experiments in which writers are randomly assigned to AI-assisted and unassisted conditions, longitudinal designs that track recall over time, and analyses of reference accuracy in both groups. Research should also examine whether particular patterns of AI use, such as using it for critique rather than composition, protect comprehension.

6. LIMITATIONS

The study has several limitations. First, the sample is small, comprising twenty papers, and purposively selected, so the proportions reported (nine of ten, eight of ten, five of ten) are descriptive and cannot be generalised statistically. Second, classification depended on AI-detection tools whose accuracy is contested (Weber-Wulff et al., 2023; Liang et al., 2023). A paper's detected AI percentage is not a direct measure of how its author used AI, and some misclassification is possible in both groups. Third, authors were not asked to specify their use of AI in detail, so we could not distinguish between, for example, the use of AI for language polishing and its use for conceptual synthesis. Fourth, recall is affected by many factors besides AI use, including the time elapsed since publication, the number of co-authors, and the author's language and interview comfort. We attempted to limit these through the telephone format and the comparison group, but cannot rule them out. Fifth, the venue comparison relied on the researchers' judgement of popularity and not on formal indicators, and the reference mismatches were not coded by type or compared systematically across groups. Sixth, the interviewers knew the group membership of each paper, which creates a risk of confirmation bias that triangulation among the three researchers only partly mitigates. Seventh, the interview framing, which in some cases presented the paper as

https://doi.org/10.63960/sp.2026.jtple.1.3.32

relevant to our study, involved an incomplete disclosure of our purpose. Finally, the influences on authors' willingness to participate may have shaped the sample, although the replacement procedure ensured that the analysed cases were voluntary. Ninth, responses were assessed through the researchers' judgement without a formal coding rubric or independent coding, and interview details such as duration and recording arrangements are not reported here. Tenth, no question required authors to apply a concept to a new situation, so unaided recall could not be fully separated from conceptual comprehension. Finally, alternative explanations such as time since publication, number of co-authors, career stage, actual contribution, language comfort, and paper complexity were not controlled. For all these reasons, the findings should be treated as exploratory and as grounds for more rigorous inquiry.

7. CONCLUSION

This study began with a suspicion that AI mediation might sever the link between an author and the text that bears their name, not because the text can be detected but because the author may not know it. In a small but cross-disciplinary sample, the suspicion found tentative support. Authors of AI-mediated papers were, in nine of ten cases, markedly disconnected from their papers' concepts, terminology, sources, methods, and contribution, even as the papers themselves were fluent, coherent, and strongly argued. Authors who wrote without AI, in eight of ten cases, answered with confidence and recalled the practice of research and publication. The AI-mediated papers also appeared in less popular journals, and half displayed mismatches in their references. These are associations in a small purposive sample and do not establish that AI use caused the differences.

We have proposed borrowed authorship as a tentative name for the condition these findings describe: ownership of a text without the understanding that ordinarily attends it. The concept does not accuse individuals of dishonesty. It identifies a structural feature of a scholarly economy in which credit attaches to texts, tools can produce excellent texts cheaply, and the cognitive labour that once linked the two is optional. In such an economy the author-function persists while the author's knowledge may not.

The remedy is not to abandon AI, which can assist reflective researchers in many legitimate ways, nor to intensify surveillance by detection software that cannot be trusted. It may be to restore comprehension as a criterion of authorship. Editors, institutions, and educators can do this by asking authors to explain, defend, and situate their work, and by rewarding the practices of close reading and interpretation through which the lived experience of research is acquired. Writing, as Plato's king warned, can give the appearance of wisdom without its reality. The task for scholarship in the age of generative AI is to ensure that, when an author signs a text, someone is there who knows what it says.

DECLARATIONS

Author(s) Contribution

Conceptualization: Dr. Ghulam Mohammad Khan; Data curation: Dr. Ghulam Mohammad Khan; Formal analysis: Dr. Ghulam Mohammad Khan; Investigation: Dr. Manzoor Ahmad Parey; Methodology: Dr. Manzoor Ahmad Parey; Resources: Dr. Mubashir Hamid Mir; Software: Dr. Mubashir Hamid Mir; Supervision: Mr. Irfan Ahmad Wagay; Writing – original draft: Dr. Ghulam Mohammad Khan; Writing – review & editing: Mr. Irfan Ahmad Wagay. All authors have read and agreed to the published version of the manuscript.

Uses of AI Acknowledgement

The authors used the AI tool (Claude) for grammatical coherence.

Funding Information

The research didn't receive any grant from any funding agency.

Availability of Data and Materials

The data supporting this study are available upon request.

Competing Interests

The authors declare no competing interests.

Clinical Trial Registration (if applicable)

Not Applicable

Human Ethics and Consent to Participate

The study did not involve any clinical interventions or experiments requiring formal ethical approval.

REFERENCES

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, *13*(3), 219–235. https://doi.org/10.1177/1088868309341564

Asagar, M. S. (2026). Navigating the future. *Journal of Interdisciplinary Studies in Education*, *15*(1), 323–344. https://doi.org/10.32674/cb7h5r51

Barthes, R. (1977). The death of the author. In *Image–music–text* (S. Heath, Trans., pp. 142–148). Fontana. (Original work published 1967)

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots. In

ACM Conference on Fairness, Accountability, and Transparency (FAccT) (pp. 610–623). https://doi.org/10.1145/3442188.3445922

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world* (pp. 56–64). Worth.

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

Carr, N. (2010). *The shallows: What the internet is doing to our brains*. W. W. Norton.

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive engagement to active learning Outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Connell, R. (2020). Southern theory. In *Routledge eBooks*. https://doi.org/10.4324/9781003117346

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/s0022-5371(72)80001-x

Derrida, J. (1981). Plato's pharmacy. In *Dissemination* (B. Johnson, Trans., pp. 61–171). University of Chicago Press.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066x.34.10.906

Foucault, M. (1977). What is an author? In D. F. Bouchard (Ed.), *Language, counter-memory, practice: Selected essays and interviews* (pp. 113–138). Cornell University Press. (Original work published 1969)

Gerlich, M. (2025). AI Tools in Society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, *15*(1), 6. https://doi.org/10.3390/soc15010006

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, *19*(1), 121–127. https://doi.org/10.1136/amiajnl-2011-000089

Hicks, M. T., Humphries, J., & Slater, J. (2024). ChatGPT is bullshit. *Ethics and Information Technology*, *26*(2). https://doi.org/10.1007/s10676-024-09775-5

International Committee of Medical Journal Editors. (2024). *Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals*. https://www.icmje.org/recommendations/

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, *55*(12), 1–38. https://doi.org/10.1145/3571730

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, 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. https://doi.org/10.1016/j.lindif.2023.102274

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X., 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. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2506.08872

Lee, H., 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 *CHI Conference on Human Factors in Computing Systems* (pp. 1–22). https://doi.org/10.1145/3706598.3713778

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, *4*(7), 100779. https://doi.org/10.1016/j.patter.2023.100779

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

Ong, W. J. (1982). *Orality and literacy: The technologizing of the word*. Methuen.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and Bias in Human Use of Automation: an attentional integration. *Human Factors the Journal of the Human Factors and Ergonomics Society*, *52*(3), 381–410. https://doi.org/10.1177/0018720810376055

Plato. (1995). *Phaedrus* (A. Nehamas & P. Woodruff, Trans.). Hackett.

Polanyi, M. (1966). *The tacit dimension*. Doubleday.

Ricoeur, P. (1981). *Hermeneutics and the human sciences* (J. B. Thompson, Ed. & Trans.). Cambridge University Press.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, *20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Rosenblatt, L. M. (1978). *The reader, the text, the poem: The transactional theory of the literary work*. Southern Illinois University Press.

Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: an illusion of explanatory depth. *Cognitive Science*, *26*(5), 521–562. https://doi.org/10.1207/s15516709cog2605_1

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology Human Learning & Memory*, *4*(6), 592–604. https://doi.org/10.1037/0278-7393.4.6.592

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. https://doi.org/10.1126/science.1207745

Stiegler, B. (1998). *Technics and time, 1: The fault of Epimetheus* (R. Beardsworth & G. Collins, Trans.). Stanford University Press.

Stokel-Walker, C. (2023). ChatGPT listed as author on research papers: many scientists disapprove. *Nature*,

613(7945), 620–621. https://doi.org/10.1038/d41586-023-00107-z

Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science*, *379*(6630), 313. https://doi.org/10.1126/science.adg7879

Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. *Scientific Reports*, *13*(1), 14045. https://doi.org/10.1038/s41598-023-41032-5

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P., & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, *19*(1). https://doi.org/10.1007/s40979-023-00146-z

Wegner, D. M. (1987). Transactive Memory: a contemporary analysis of the group mind. In *Springer eBooks* (pp. 185–208). https://doi.org/10.1007/978-1-4612-4634-3_9

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 do not necessarily reflect the views of Synergy Publication or the editor(s) of the journal. While reasonable editorial and peer-review procedures are followed, Synergy Publication and the editor(s) do not endorse any products, methods, or claims described in the content.

Author Bio:

1. Dr. Ghulam Mohammad Khan is a writer, literary scholar, and academic from Kashmir. His writing moves between fiction, personal essay, and literary criticism, exploring memory, history, identity, power, and the intimate lives shaped by contested histories. His work is particularly interested in how ordinary lives and forgotten stories become repositories of larger historical and political truths. Khan holds a PhD in English from the Central University of Haryana, where he researched narratology and memory in postcolonial literature. He is the author of Anecdote, History and Kashmir: And Other Essays (2025), a collection that brings together literary criticism, history, memory, and personal experience to question received narratives of Kashmir. His fiction and essays have appeared in The Rumpus, Indian Literature (Sahitya Akademi), India Forum, KITAAB, Out of Print, The Wire, Madras Courier, Muse India, Kashmir Lit, Gulmohur Quarterly, and other literary publications. His long-form essay was published by The Rumpus in August 2026. He is presently an Assistant Professor of English at HKM Government Degree College, Bandipora, Kashmir.

2. Dr. Manzoor Ahmad Parey is an Assistant Professor of Education in the Department of Higher Education, Government of Jammu & Kashmir, presently serving at Government Degree College, Bandipora. He completed his Ph.D. in Educational Studies from Jamia Millia Islamia, New Delhi, in 2026. His doctoral research examines schooling and educational realities in Kashmir. His academic interests include school education, critical pedagogy, educational policy and reform and emerging questions concerning AI and education.

3. Dr. Mubashir Hamid Mir presently works as an Assistant Professor in Physics at HKM Govt. Degree College Bandipora, Jammu and Kashmir. He completed his MSC and Ph.D in Physcis from the University of Kashmir.

4. Mr. Irfan Ahmad Wagay presently works as an Assistant Professor in Kashmiri at HKM Govt. Degree College Bandipora, Jammu and Kashmir. He completed his MA and MPhil in Kashmiri from the University of Kashmir.