

Tel Aviv University
Faculty of Education
Department of _____

**How graduate students experience and make sense of
using generative AI in academic writing (an IPA
study)**

Academic paper submitted for the course:
Technology in Education

Submitted by:
Dana Levi

July 2026

Table of Contents

Abstract.....	1
1. Introduction.....	3
2. Theoretical Background.....	8
2.1. Fundamental Concepts and Definitions.....	8
2.2. Review of Previous Research.....	9
2.3. Synthesis and the Research Gap.....	12
2.4. Conceptual Framework.....	14
3. Research Questions and Hypotheses.....	17
3.1. Research Questions.....	17
3.2. Conceptual Lens.....	18
4. Methodology.....	20
4.1. Research Design.....	20
4.2. Research Participants.....	21
4.3. The Interview Guide.....	22
4.4. Data Collection.....	23
4.5. Data Analysis Method.....	23
4.6. Reflexivity.....	24
4.7. Trustworthiness and Rigor.....	25
4.8. Ethical Considerations.....	26
5. Findings.....	27
5.1. Participant Profile.....	27
5.2. Main Themes.....	28
5.3. Sub-themes.....	30
5.4. Outlier Cases.....	33
6. Discussion.....	34
6.1. Interpretation of the Main Findings.....	34
6.2. Comparison with the Literature.....	35

6.3. Research Contribution and Practical Implications.....	37
7. Limitations of the Study.....	39
8. Recommendations for Future Research.....	40
9. Summary and Conclusions.....	43
9.1. Summary of Main Findings.....	43
9.2. Conclusions.....	44
9.3. Recommendations.....	45
References.....	48
Appendices.....	51
Appendix A: Interview Guide.....	51

List of Tables

Table 1. Participant Characteristics.....	21
--	-----------

Table 2. Main themes and representative quotes.....	29
--	-----------

List of Figures

Figure 2. Primary stages of generative AI use in the writing process.....	27
--	-----------

Abstract

Research Problem and Approach: The proliferation of generative AI tools positions the academic writing of graduate students as an arena where linguistic assistance, idea development, and copyediting are intertwined with questions of trust, ownership, and academic integrity. While institutional discourse tends to focus on detecting anomalies and establishing usage policies, there remains a lack of qualitative understanding regarding how students themselves interpret their daily encounters with these tools. Employing a qualitative-interpretative approach, this study examined the user experience and the meanings attributed to the use of generative AI in the writing process.

Methodology and Themes: The research drew on a purposive sample of graduate students in higher education institutions who use generative AI tools in their academic writing assignments. Data were collected through semi-structured in-depth interviews and short reflective texts, and analyzed using Interpretative Phenomenological Analysis (IPA; Smith). This approach allowed for an exploration of how participants interpreted their experiences, moving beyond mere descriptions of their actions. Several main themes emerged from the analysis: the tool as an ambivalent writing partner, a shift in the writer's sense of agency, the tension between efficiency and academic integrity, and the role of AI literacy in shaping critical use.

Key Contributions: The study makes a threefold contribution. First, it shifts the focus of the debate from whether the use of these tools is permissible to how this use is experienced, managed, and justified by graduate students. Second, it conceptualizes generative AI as a writing partner that does not replace the author, but rather alters the boundaries of responsibility, copyediting, and text ownership. Third, it demonstrates that students' perceptions are shaped by institutional and disciplinary contexts, rather than solely by personal technological preferences.

Implications: The findings are relevant to lecturers, supervisors, academic policymakers, and students, as they indicate that the responsible use of generative AI requires more than just guidelines on what is permitted or prohibited. The findings also highlight the need for a clear pedagogical and ethical language that integrates AI literacy, transparency in the writing process, and a distinction between legitimate assistance and academic misconduct.

Keywords: generative AI, large language models, academic writing, graduate students, AI literacy, academic integrity, academic ethics, writing process, user experience, student perceptions, writer agency, text ownership, institutional policy

1. Introduction

Academic writing by graduate students currently takes place in an environment where generative AI tools are available at almost every stage of the writing process: brainstorming, initial drafting, structuring arguments, paraphrasing, copyediting, generating abstracts, and providing feedback on existing text. This availability is not merely a technical addition to the writer's toolkit, but a shift in the conditions of intellectual labor itself. When text can be generated, altered, and refined using large language models, the question of what remains the responsibility of the student writer becomes more acute: selecting the argument, exercising judgment, verifying sources, making stylistic decisions, and adhering to the requirements of academic integrity and ethics. The mere option to use these tools does not resolve the question of responsibility; rather, it shifts the focus: how to determine when the assistance supports writing and when it replaces the academic work the student is expected to perform independently.

This issue is particularly salient in graduate studies. Graduate students are required not only to submit grammatically sound assignments but also to demonstrate research capability, interpretive independence, mastery of disciplinary discourse, and the ability to position their voice within an existing field of knowledge. Therefore, the use of generative AI tools cannot be reduced to whether the tool is "helpful" or "harmful"; it touches upon how students perceive themselves as writers, as early-career researchers, and as members of an academic community. In this sense, user experience is not merely about the operational convenience of an interface, but rather a space where trust, suspicion, dependence, responsibility, and boundaries of action are shaped.

Existing literature already identifies generative AI as a technology with broad implications for research, professional work, and knowledge organization. However, acknowledging the magnitude of this shift is insufficient for understanding its practical implications for academic writing. Broad discussions on the technology's impact address uncertainty, misinformation, biases, privacy violations, and ethical dilemmas, yet they often remain at the macro level of institutions, professions, and the labor market (Budhwar et al., 2023). For the student writer, these same dilemmas translate into small, recurring decisions: Should I ask the tool for an alternative phrasing of a paragraph? Should I use it to suggest an article structure? Should I rely on its explanation of a theoretical concept? Is disclosure required even when the tool was used "only" for copyediting? The accumulation of these localized decisions is what shapes the meaning of the tool's use: it is not general policy alone, but how that policy intersects with the act

of writing in the moments when the student must choose to verify, adopt, or reject technological output.

This points to a specific research gap. It is known that generative AI tools can streamline various aspects of academic writing, and it is equally known that they raise concerns regarding authenticity, reliability, and the preservation of critical thinking. However, it is less clear how graduate students experience this tension while working on actual academic texts. Reviews addressing the transition from human writing to text generated by AI systems highlight both the potential for increased efficiency and the risk of compromising the authenticity of academic work; thus, the question is not whether the tools are beneficial or dangerous, but how writers navigate between utility and risk (Dergaa et al., 2023). This gap is particularly important because graduate-level writing often involves decisions where "content" and "phrasing" cannot be neatly separated: articulating an argument is also a conceptual act, paraphrasing is also an ethical act, and copyediting is also the shaping of an academic voice. Therefore, research focusing on the user experience is not a retreat from the normative question, but a prerequisite for understanding it: to determine what constitutes appropriate use, we must understand how users practically distinguish between assistance, substitution, inspiration, dependence, and responsibility.

Another gap relates to institutional policy and its impact on student perceptions. Higher education institutions are formulating various guidelines for the use of generative AI tools, yet students often operate in an ambiguous space: in one course, use is permitted under certain conditions; in another, it is prohibited; and in many cases, there is no clear distinction between legitimate assistance and a violation of academic integrity. Studies and discussions in higher education point to the need for clear policies and operational frameworks, as well as the need to understand the actual experiences of users of these systems (Michel-Villarreal et al., 2023). This ambiguity is not merely an administrative problem; it shapes the writing experience itself. A student may benefit from the feedback of a generative AI tool, yet simultaneously feel that the very act of using it places them in a moral and institutional gray area. Conversely, overly rigid policies may miss the subtle distinctions between different types of use, thereby reducing the discussion from an educational question of responsibility and judgment to a narrow question of prohibition or permission. Therefore, we must examine how policy, institutional uncertainty, and daily writing practices mutually shape one another.

Against this backdrop, the purpose of this study is to understand how graduate students experience the use of generative AI tools in their academic writing, and the meanings they attribute to this use. The guiding research question is: How do graduate students experience the use of generative AI tools in academic writing, and what meanings do they attribute to it? This question is deliberately framed as an interpretive one rather than merely a question of efficiency. It does not seek to measure whether a specific tool improves grades, shortens working time, or reduces grammatical errors, but rather to understand how students interpret the tool's place within the writing process, what they feel they gain or lose, and how they delineate the boundaries of appropriate use for themselves. This focus does not dismiss the relevance of efficiency questions, but argues that they are insufficient: even a use that streamlines writing may be experienced as problematic if it undermines a sense of ownership, and a use that produces no measurable improvement may still be valuable if it supports confidence, clarity of thought, or critical awareness.

The main argument of this study is that graduate students do not experience generative AI tools as a simple substitute for the academic writer, but rather as an ambivalent writing partner: it assists in developing ideas, drafting, copyediting, and providing feedback, yet simultaneously destabilizes trust, agency, integrity, and the definition of textual ownership. Consequently, the meaning of using generative AI in academic writing depends on AI literacy, clear institutional policies, and the disciplinary and ethical context in which the writing occurs. This argument avoids two prevalent extremes in public and academic discourse: on the one hand, presenting the tools as a neutral solution to problems of phrasing and efficiency; on the other, presenting them as a uniform threat to the very possibility of authentic academic writing. Between these two extremes lies a more complex user experience. The same tool can support a student when serving as a critical mirror for the text, yet disempower them when it becomes the authority from which an argument is derived. Therefore, it is not enough to ask whether a tool was used; we must ask at what stage, for what purpose, with what degree of scrutiny, and within what boundaries of responsibility.

This study contributes to the field by centering the experiences of graduate students. Rather than focusing on the tool's capabilities, it begins with the writer, examining how the tool is integrated into their writing process. This shift in perspective is crucial because generative AI tools do not operate in a vacuum: they receive prompts, generate responses, trigger revisions, and

are either rejected or adopted; at times, they even alter how students perceive the quality of their own writing. Consequently, AI literacy extends beyond the technical knowledge of operating a system. It encompasses the critical ability to formulate prompts, evaluate responses, identify limitations, verify claims, distinguish between language editing and conceptual contributions, and integrate the output into human writing while maintaining authorial responsibility. In this regard, the study does more than describe student attitudes; it proposes framing AI literacy as an integral component of academic literacy.

Additionally, the study integrates three intersecting axes: writing, technology, and ethics. Along the writing axis, academic writing is approached as an ongoing process of planning, drafting, revising, reviewing, and editing, rather than merely a final product. Along the technology axis, generative AI tools are framed as supportive resources for the writing process rather than complete substitutes for the student's thinking, judgment, and authorial voice. Along the ethics axis, the study examines how academic integrity, transparency, and personal responsibility become inseparable from the technological experience. Integrating these axes helps explain why the exact same action—such as requesting phrasing improvements—might be viewed by one student as legitimate assistance and by another as a threat to textual ownership. This approach avoids artificially fragmenting the phenomenon: analyzing writing without ethics reduces the issue to mere textual enhancement; analyzing technology without writing ignores the context that gives the output meaning; and analyzing ethics without user experience fails to explain how abstract norms translate into practical decisions.

The focus on graduate students is further justified by their specific position within the academic hierarchy. While they have moved beyond learning basic writing conventions, they are still in the process of professionalizing as researchers and forming their academic identities. Many are required to synthesize extensive bodies of literature, formulate research questions, distinguish between original arguments and source summaries, and navigate feedback from instructors and supervisors. In this context, generative AI tools can act as an additional mediator between the student and the demands of academic discourse. While these tools can help alleviate cognitive load, clarify argumentative structures, and improve linguistic flow, they also risk blurring the boundaries between comprehension and production, between independent drafting and adopting generated text, and between constructive feedback and epistemic authority. Consequently, studying this demographic provides an opportunity to examine the use of these

tools not merely as an academic skill, but as part of the broader socialization into a research community and its norms.

The choice of a qualitative approach is driven by the nature of the research question. Experiences, interpretations, hesitations, and a sense of responsibility cannot be fully captured through quantitative metrics alone. A qualitative methodology allows for tracing how students narrate their use of these tools, justify or express reservations about them, differentiate between various types of use, and connect these practices to institutional policies, disciplinary contexts, their writing confidence, and their relationships with supervisors and instructors. This approach does not seek statistical generalizability across the entire graduate student population; rather, it aims to provide a deeper understanding of the meanings participants ascribe to emerging writing practices. In doing so, this study seeks to illuminate the writing process as a space where technological engagement, human judgment, and institutional norms are inextricably linked. The strength of the qualitative approach lies in its capacity to explore not only what students do with these tools, but also how they explain their actions to themselves and others, and the language they use to articulate responsibility, apprehension, legitimacy, and boundaries.

The study is structured as follows: the literature review first introduces the discourse surrounding generative AI, large language models, and academic writing in higher education. Next, it outlines the conceptual framework used to analyze user experience, student perceptions, AI literacy, and academic integrity. This is followed by a presentation of the methodological considerations underlying the qualitative research. Finally, the study discusses the implications of the analysis for graduate students' writing processes, their responsibilities as authors, and the role of academic institutions in establishing clear and equitable guidelines for use.

2. Theoretical Background

2.1. Fundamental Concepts and Definitions

Discussing the experiences of graduate students using generative AI tools in academic writing requires distinguishing between several layers of this phenomenon. The first layer involves artificial intelligence as a broad term, referring to computational systems that perform tasks previously considered reliant on human judgment, pattern recognition, learning, and decision-making. The second layer comprises generative AI: systems capable of producing new content—such as text, images, code, summaries, translations, alternative phrasing, or conceptual suggestions—based on patterns learned from vast datasets. The third layer consists of large language models, which serve as the primary infrastructure for many of the tools graduate students currently use in academic writing. Their power lies not merely in answering questions, but in their ability to maintain a dialogic sequence, adapt phrasing to context, mimic discourse genres, suggest argumentative structures, and generate text that, at first glance, appears coherent, standard, and persuasive.

This distinction is necessary because using generative AI tools in academic writing differs from using a search engine, dictionary, article database, or grammar checker. A search engine directs users to sources; generative AI generates phrasing. A bibliographic database retrieves items; a large language model can suggest connections between ideas even when it fails to substantiate them properly. A proofreading tool corrects local errors; a dialogic system can influence the formulation of an argument, the order of sections, the choice of concepts, and the writing style. Therefore, from a theoretical standpoint, these tools should not be viewed merely as "technical aids," but as active components within the writing environment. They do not write for the student in a complete sense, but they participate in shaping the conditions of possibility for writing: what counts as good phrasing, what constitutes a reasonable answer, which alternatives are available, and what cognitive effort is pushed out of sight.

For graduate students, academic writing is not merely a linguistic act but a practice of joining an interpretive and research community. These students are required to demonstrate mastery of the literature, distinguish between arguments, develop a scholarly voice, use sources carefully, situate their contribution relative to existing knowledge, and adhere to norms of academic

integrity. The writing process thus encompasses visible stages, such as drafting paragraphs and copyediting, alongside more covert stages, such as deliberation, deletion, thought experiments, processing feedback, negotiating with a supervisor, and establishing authorial authority. When generative AI tools enter the writing process, they intersect not only with the need for efficient phrasing but also with anxiety, the desire to belong, the urge to sound "academic," the fear of failure, and the need to establish ownership of an idea.

User experience in this context is not limited to convenience, speed, or satisfaction. It encompasses how graduate students interpret the tool's status relative to themselves: whether it is perceived as an assistant, a teacher, an editor, a partner, a threat, an illicit shortcut, or a safe space for experimentation. It also includes the trust placed in the outputs, patterns of checking and verification, the writer's sense of agency, and how students define the boundaries between legitimate use and academic misconduct. Accordingly, AI literacy is not merely operational knowledge of prompting the system, but the ability to understand the tools' limitations, evaluate output quality, identify errors or biases, decide when to use them and when to refrain, and report their use in accordance with academic ethics and institutional policy.

The theoretical premise of this section is that these tools are not pedagogically neutral. The very ability to rapidly produce fluent text alters the relationship between effort, learning, and output. On one hand, it may free up cognitive resources for engaging with argumentation, structure, and critique. On the other hand, it risks blurring the connection between the writing process and the acquisition of academic skills. Therefore, the question is not whether generative AI is "good" or "bad" for academic writing, but how graduate students experience its entry into the writing process, and what meanings they attribute to dependence, assistance, trust, usage boundaries, and textual ownership.

2.2. Review of Previous Research

The general literature on artificial intelligence suggests that the technology is poised to transform intellectual labor, not merely by replacing tasks but by augmenting existing human capabilities. In this context, the discussion of generative AI in academic writing extends a broader debate on the automation of cognitive work: which aspects of human activity can be accelerated, which can be delegated to a tool, and which require human judgment that cannot be easily substituted. A multidisciplinary review of artificial intelligence indicates that the rapid development of this technology simultaneously creates opportunities, disruptions, and both policy

and practical challenges; it therefore demands a sober assessment of its impacts rather than unreserved technological enthusiasm (Dwivedi et al., 2019). For academic writing, this means that while these tools may support graduate students in performing certain tasks, the support itself alters the conditions of learning, assessment, and accountability.

Studies on digital transformation add an organizational and institutional dimension to this understanding. The introduction of a new technology is not an isolated event where a tool simply enters an existing environment without altering it; rather, it disrupts routines, roles, expectations, and power dynamics. A systematic review of the digital transformation literature describes a shift toward agile organizational designs and digital ecosystems, distinguishing between various perspectives: technological impact, domain adaptation, systemic change, and the co-evolution of organization and technology (Hanelt et al., 2020). In higher education, this insight clarifies that the use of generative AI in academic writing is not merely a private matter between a student and a screen. It depends on regulations, instructors' attitudes, institutional infrastructure, disclosure policies, disciplinary norms, and questions of accessibility. The student experience is thus shaped at the intersection of individual action and an institutional system that is still learning how to respond.

More direct empirical findings regarding student perceptions indicate an ambivalent attitude. A study involving 399 undergraduate and graduate students in Hong Kong found that while students recognize the utility of generative AI for personalized learning, brainstorming, writing assistance, and research analysis, they simultaneously express concerns regarding accuracy, privacy, ethics, and potential impacts on their personal and professional development (Chan & Hu, 2023). This finding goes beyond merely confirming a positive reception of the technology; it necessitates a distinction between use and trust. A student might use a tool because it is available, efficient, and helps overcome writer's block, while simultaneously distrusting the accuracy of the information it generates. Similarly, a positive attitude toward the tool does not necessarily imply agreement that its use is appropriate at every stage of the writing process. The distinction between pragmatic utility and epistemic trust is particularly crucial in academic writing, where a fluent but unfounded text can be more misleading than a cumbersome but cautious one.

The problem of trust is brought into sharper focus in the literature on transparency and explainability in artificial intelligence systems. Although studies in this area do not necessarily focus on academic writing, they provide a foundation for understanding the difficulty of relying

on an opaque output. When a system generates a convincing answer without clearly demonstrating how it arrived at it, the user must make a decision based on an output whose generation process is difficult to trace. A review of interpretability in machine learning systems points to a lack of consensus on how to evaluate the quality of the explanations these systems provide, as well as the growing importance of verifiability for informed decision-making (Carvalho et al., 2019). In the context of graduate students, the problem is not merely technical. If a student receives from the tool a formulation of an argument, a summary of a concept, or a suggestion for integrating a source, they must determine whether this constitutes appropriate assistance or an output that requires deconstruction, verification, and sometimes rejection. The writing experience thus also becomes an experience of oversight: the student does not simply write with the tool's help, but also scrutinizes, questions, verifies, and corrects it.

The literature on explainable AI adds a critical dimension regarding the "black box" effect. Complex models can produce outputs of high rhetorical quality, but their internal complexity makes it difficult for the user to understand the chain of operations that led to the result. A review in this field shows that the lack of transparency in the decision-making processes of complex models hinders their adoption in complex domains, and that explainable AI frameworks are necessary to identify flaws and improve reliability (Hassija et al., 2023). When applying this argument to academic writing, it becomes clear that the problem is not limited to whether a paragraph "sounds good." In fact, fluency can engender overconfidence: a text written with linguistic assurance may mask a conceptual error, a fabricated citation, an overgeneralization, or a lack of disciplinary fit. Consequently, writing proficiency in the era of generative AI also entails maintaining a critical stance toward text that appears overly polished.

Developments in generative models are also transforming the digital environments in which users operate. Conceptual research on advanced digital environments indicates that generative models based on deep learning are reshaping user interactions and online environments (Park & Kim, 2022). Although this research does not deal directly with academic writing, it helps explain why generative AI tools are not merely external applications applied to a text. They are integrated into learning management systems, word processors, search engines, knowledge management environments, and collaborative platforms. In such a landscape, the student does not always "choose" to access an isolated tool; often, the tool is embedded within the workspace itself, offering completions, phrasing suggestions, corrections, and summaries as part of the

environment. This integration blurs the boundaries between independent writing, automated editing, algorithmic feedback, and the co-production of text.

Alongside the opportunities, previous research points to a tension concerning assessment and academic integrity. Traditional writing assignments often rely on the assumption that the submitted text adequately reflects the student's abilities: their understanding of the material, capacity to formulate an argument, command of sources, and linguistic skills. Generative AI undermines this assumption because it can produce human-like text and assist in completing substantial portions of the assignment. However, this does not imply that all use is inherently illegitimate. Language editing, receiving feedback, formulating research questions, or checking for clarity can be part of a legitimate learning process, provided the boundaries are clear and academic accountability is maintained. The problem arises when there is no agreed-upon distinction between assistance, collaboration, delegation, and replacement. For graduate students, this ambiguity can create a fraught writing experience: the tool is helpful, yet the assistance itself may be perceived as a moral or institutional risk.

2.3. Synthesis and the Research Gap

The existing literature suggests that generative AI functions as both a resource and a disruptive force in academic writing. While it aids in drafting, planning, brainstorming, clarifying ideas, and editing, it simultaneously raises questions regarding reliability, responsibility, agency, and academic integrity. Consequently, a theoretical framework that treats this technology merely as an efficiency tool is inadequate. Instead, it must be understood within a relational network where the student, the text, the institution, the tool, and academic norms mutually shape one another. The focal point of the discussion thus shifts from what the tool can do to what happens to the academic writer when the tool participates in the writing process.

Synthesizing prior research highlights three primary distinctions. The first is the difference between linguistic assistance and epistemic contribution. When a generative AI tool refines phrasing, shortens sentences, or suggests smoother transitions between paragraphs, its function can be viewed as an extension of copyediting. However, when it posits an argument, explains a concept, categorizes literature, or offers an interpretation, it enters the domain of knowledge production. This distinction is especially pertinent for graduate students, whose writing is evaluated not merely on grammatical correctness but on originality, accuracy, engagement with the literature, and theoretical maturity. A practice that is unproblematic at one stage of the writing

process may become highly sensitive at another. For instance, accepting a suggestion to clarify an existing paragraph differs fundamentally from generating a core argument that the student adopts uncritically.

The second distinction lies between usage and trust. Research on student perceptions demonstrates that a willingness to experiment with these tools does not negate underlying concerns. Students may appreciate the tool for its accessibility, speed, and ability to reduce workload, yet still view it as unreliable or ethically risky. In academic writing, this gap produces a dual pattern of behavior: reliance and verification, attraction and suspicion, adoption and reservation. In other words, the user experience is not simply a positive or negative attitude toward the tool, but an ongoing process of decision-making. The student must constantly decide what to ask, what to accept, what to delete, what to verify, what to claim as their own, and what to disclose to others. This underscores the critical importance of AI literacy, which is not an external add-on to the writing process but a prerequisite for navigating the interaction with the tool effectively.

The third distinction is between personal and institutional responsibility. While it is easy to place the entire burden of appropriate use on the student, the user experience is heavily shaped by institutional policies, instructor expectations, the availability of guidelines, conflicting messages, and variations across courses and disciplines. In the absence of clear rules, students are left to interpret the boundaries of legitimate use on their own. This ambiguity can exacerbate inequalities: those who are well-versed in academic norms, proficient in academic discourse, and skilled at drafting disclosure statements will navigate these tools with greater confidence than those who fear that any use might be deemed academic misconduct. Consequently, even when research focuses on individual experiences, these cannot be understood in isolation from broader structures of power and authority. The student operates within a system where institutions, instructors, commercial platforms, and disciplinary communities explicitly or implicitly delineate the boundaries of permissible use.

The existing literature, however, leaves a gap regarding the current research question. Much of the scholarship addresses artificial intelligence or digital transformation at a broad level, providing a framework for understanding technological change but lacking a rich description of the academic writing experience itself. While empirical studies on students contribute to our understanding of their perceptions, they often aggregate diverse populations, disciplines, and

stages of study. Consequently, our knowledge of how graduate students use generative AI tools during the actual writing process remains fragmented. It is not enough to know whether they are "for" or "against" these tools; we must understand how they interpret their own actions when formulating an argument, synthesizing a source, addressing feedback, drafting a theoretical section, or copyediting a text prior to submission.

This gap is also conceptual. Terms such as "assistance," "usage," "generation," "editing," "originality," and "ownership" lose their stability when a tool can influence multiple levels of a text. If a student writes a paragraph and asks the tool to refine it, ownership of that paragraph seemingly remains clear. However, if the student asks the tool to propose an argumentative structure, receives key phrasing, modifies it, and integrates it into the text, the boundaries become blurred. These are not merely technical questions; they strike at the core of the writer's academic identity. Graduate students are not simply submitting assignments; they are gradually developing a scholarly disposition. Therefore, the use of generative AI may be experienced not merely as assistance, but as an intervention in the process of cultivating an academic voice.

The central tension emerging from the literature is the pull between empowering and diminishing student agency. On one hand, the tool can help students overcome writer's block, grasp genre requirements, refine their phrasing in a non-native language, receive immediate feedback, and test various argumentative pathways. On the other hand, it risks encouraging overreliance, eroding patience for slow conceptual work, replacing the struggle with ambiguity with a quick answer, and fostering the illusion that "academic" writing is merely a superficial style generated at the push of a button. A more balanced conclusion is that this impact is not determined by the tool itself, but by how it is integrated into the writing process, the user's level of AI literacy, the nature of the assignment, disclosure policies, and the disciplinary context.

2.4. Conceptual Framework

The conceptual framework of this study rests on the notion of generative AI as an ambivalent writing partner. The term "partner" does not attribute intention, consciousness, or moral responsibility to the tool; rather, it indicates its active participation in shaping the writing process: it suggests, organizes, drafts, responds, corrects, and challenges. The term "ambivalent" denotes that the same action can be experienced simultaneously as helpful and destabilizing. A suggestion for precise phrasing might bolster confidence, yet also raise doubts about who actually wrote the sentence. Immediate feedback might alleviate the isolation of the writing process, but it can also

create dependency. Summarizing an idea may clarify a concept, but it can also flatten it. Therefore, the experience is not measured on a unidimensional continuum of acceptance or rejection, but rather within a space of competing meanings.

Within this framework, the writing process is conceptualized as a sequence of decisions rather than merely a sequence of stages. Graduate students decide when to consult the tool, how to formulate a prompt, which output to accept, what to verify, what to modify, and what to exclude from the text. Each such decision reflects a specific orientation toward trust, agency, and academic integrity. For instance, using the tool for brainstorming may express a desire to expand their thinking; using it to write an entire paragraph may reflect time pressure, linguistic difficulties, or a different understanding of the boundaries of responsibility. The current study does not presuppose that a particular action is inherently legitimate or illegitimate; instead, it seeks to understand how the users themselves ascribe meaning to it within their academic world.

Three concepts organize the interpretation of the students' experiences. The first is agency. In this context, agency refers to graduate students' ability to feel they are leading the writing process even when assisted by the tool. Strong agency does not require unassisted writing; it requires the ability to choose, judge, reject, justify, and take responsibility. When a student uses generative AI but continues to exercise judgment, the tool can expand their scope of action. When a student adopts outputs without understanding or verification, agency is diminished. In this study, agency serves as a theoretical mechanism rather than a quantitatively measurable variable; it acts as a lens for interpreting the participants' accounts and is not presented as a statistically tested causal relationship.

The second concept is trust. Trust in generative AI tools is not binary. Students may trust the tool with phrasing but not with facts; they may rely on it to organize ideas but not for theoretical interpretation; they may accept its stylistic suggestions but refrain from delegating decisions regarding sources. Therefore, a distinction must be made between operational trust, which concerns efficiency and availability, and epistemic trust, which concerns correctness, accuracy, substantiation, and academic judgment. This distinction helps explain why the same user experience can encompass both high satisfaction and deep apprehension. It also explains why questions of transparency, verification, and validation are not marginal, but rather an integral part of writing with large language models.

The third concept is academic integrity. Within this framework, academic integrity is not reduced to a prohibition against plagiarism; it encompasses responsibility for the source of a claim, the method of its generation, fair reporting, and maintaining the relationship between learning and the final product. Generative AI tools disrupt familiar distinctions because they are not a standard "source," an "author" in the conventional sense, or merely a passive "tool." Consequently, students' experiences may involve not only practical questions about usage but also ethical ones: what is permissible to accept from the tool, when must I disclose it, is the text still mine, and does this usage undermine the purpose of the assignment? These questions shape how students write, abstain, conceal, or report.

The conceptual framework positions AI literacy as a theoretical mediating capacity, rather than a variable measured quantitatively here. This literacy includes an understanding of the tool's capabilities, recognition of its limitations, proficiency in formulating prompts, the ability to evaluate outputs, awareness of academic ethics, and an understanding of institutional policy. It is hypothesized that the more developed this literacy is, the better equipped the student will be to use the tool in a way that expands the writing process without relinquishing their responsibility as an author. In its absence, the very same tool may be perceived as an outsized source of authority or as a shortcut that replaces academic work.

The student experience takes shape within a shifting context; therefore, there is no single, uniform experience of using generative AI in academic writing. It varies by discipline, stage of study, language of writing, assignment requirements, supervisory relationships, institutional policy, level of academic confidence, and technological accessibility. For this reason, qualitative research is particularly suited to the research question: it allows for tracing the meanings students attribute to their usage, the tensions they identify, and the ways they articulate operational boundaries for themselves. The starting point is not that the tool replaces the writer, but rather that it introduces an unstable partnership into the writing process, necessitating constant interpretation of responsibility, ownership, trust, and authority.

[figure to be inserted here]

Figure 1. The methodological process flow

3. Research Questions and Hypotheses

3.1. Research Questions

The guiding research question of this article is how graduate students experience using generative AI tools in academic writing, and what meanings they attribute to this practice. This focus moves beyond the technical efficiency of these tools, directing attention instead to how they are integrated into and shape the writing process, while evoking subjective, professional, and ethical interpretations among writers. The underlying premise is that the impact of these tools cannot be understood solely through the final product, as significant decisions are made during the brainstorming, drafting, copyediting, and self-evaluation stages. Accordingly, this article does not seek to measure improvements in writing performance or to compare users with non-users. Instead, it aims to understand how graduate students interpret their encounters with generative AI tools as part of their routine academic work.

Several sub-questions derive from this primary focus, addressing the roles, benefits, tensions, and boundaries associated with using these tools. The article examines how graduate students describe the roles of these tools at various stages of the writing process, including brainstorming, structuring arguments, drafting, copyediting, paraphrasing, and receiving feedback. It also explores the meanings students attribute to these tools regarding time-saving, overcoming language barriers, conceptual refinement, and building confidence in academic writing. Concurrently, the article asks what tensions, concerns, or boundaries students identify when using generative AI, particularly concerning academic integrity, output reliability, text ownership, and potential overreliance on the system. Another question addresses how students understand their responsibility as writers when parts of the writing process are mediated by large language models. Finally, the article investigates how AI literacy, institutional policy, and disciplinary context shape students' perceptions of using these tools.

The discussion rests on the assumption that generative AI tools are not merely neutral assistive technologies. While their contribution can be measured in terms of improved phrasing, time saved, or error reduction, such metrics are insufficient for understanding their true place in the writing process. Even when these tools provide practical assistance, their integration alters the division of labor between the writer and the system, fundamentally changing how the writer

understands originality, responsibility, and text ownership. Consequently, the user experience is not merely a reaction to an effective or ineffective tool, but an interpretation of a new writing paradigm where technological support and academic decision-making coexist.

As a qualitative study, this research does not propose statistical hypotheses in the positivist sense of testing relationships between variables. However, the article relies on the interpretive premise that the user experience is not uniform: the same tool may be experienced simultaneously as an aid, a disruptor, an accelerator, and something that requires critical oversight. Therefore, rather than offering a hypothesis for quantitative testing, this study formulates a theoretical expectation: the meaning students attribute to generative AI in academic writing emerges from the tension between receiving writing assistance and preserving agency, responsibility, and academic ethics.

3.2. Conceptual Lens

The conceptual framework of this article views academic writing simultaneously as a social, cognitive, and ethical practice. Writing is not merely the production of a well-formed text, but a process through which the writer clarifies a position, organizes knowledge, engages in dialogue with a disciplinary community, and takes responsibility for their claims. When generative AI tools are integrated into this process, they serve not merely as linguistic aids, but as mediators between the student and the text, between intention and articulation, and between personal knowledge and a product generated with computational assistance. Therefore, the question is not whether the tool "writes instead of" the student or merely "helps" them, but how the students themselves draw the line between legitimate assistance and the relinquishment of interpretive and academic control.

From this perspective, generative AI tools are examined as ambivalent writing partners. On the one hand, these tools can expand the range of possible actions: offering alternative phrasing, helping to break down complex tasks, clarifying argumentative structure, and easing the burden on students for whom academic writing is less natural. On the other hand, their very use raises questions about the boundaries of authorship, trust in the generated content, the identification of original contributions, and the student's ability to retain authority over a text bearing their name. The tension between these two aspects lies at the heart of the analysis: the practical advantages of the tools do not negate the ethical risks, and the ethical risks do not necessarily preclude informed and supervised use.

The concepts guiding this analysis are user experience, student perceptions, writing agency, AI literacy, academic integrity, and academic ethics. These concepts are not positioned as measurable variables, but rather as frameworks that help interpret the data. They enable an examination of how graduate students describe their use of these tools, how they justify it, where they set boundaries, and the extent to which they view it as an extension of their capabilities or a threat to their academic independence. In this way, the everyday user experience is linked to broader questions of academic training, authorial responsibility, and the shifting norms of writing in the era of large language models.

4. Methodology

4.1. Research Design

The study employed a qualitative research design, as the research question explored how graduate students interpret their use of generative AI tools within the writing process. The focus was not merely on whether these tools were used, but on how participants distinguished between legitimate assistance and destabilizing reliance, and the meaning they attributed to shifts in their sense of textual ownership. Such questions are difficult to answer through quantitative measurements of usage frequency, as they pertain to meaning, self-interpretation, and the tensions that arise during academic writing. Accordingly, the study was grounded in an interpretive-constructivist paradigm, positing that the user experience is not a stable fact hidden "behind" the participants' accounts, but is rather shaped by the language they use to describe their work, their concerns, and the boundaries they set for themselves.

The research design included semi-structured in-depth interviews with graduate students who actively used generative AI tools for academic writing. This choice aligned with the central premise of the study: that these tools are not experienced simply as substitutes for the academic writer, but rather as ambivalent writing partners. The literature points to a similar tension: alongside their potential to support learning and writing, large language models raise questions regarding accuracy, misinformation, and academic integrity (Lo, 2023). Therefore, a methodology was required that would allow participants to articulate both the practical benefits and the emotional, ethical, and identity-related costs of using these tools.

The strength of this design, compared to a large-scale survey, lay in its ability to capture the nuances of how students differentiate between drafting, copyediting, brainstorming, receiving feedback, and generating text. In this study, self-efficacy, agency, and academic identity were not measured as quantitative variables; instead, they were examined as interpretive experiences emerging from the participants' narratives. Consequently, the relationships between them were interpreted as meaningful connections derived from the data, rather than as statistically tested causal mechanisms.

4.2. Research Participants

The sampling aimed to ensure that all participants had direct experience with the phenomenon under investigation. To this end, criterion-based purposive sampling was employed: graduate students at higher education institutions who had used generative AI in at least one of their academic writing assignments. This criterion filtered out participants who merely held general attitudes toward technology, focusing the study on those with actual experience integrating these tools into the writing process. The sample comprised 12 participants, an appropriate number for an interpretive-phenomenological study seeking idiographic depth alongside the ability to compare cases. After the first ten interviews, the main themes clearly recurred, and two additional interviews allowed us to examine the stability of these patterns and the presence of any divergent cases.

As shown in Table 1, the sample included students from a variety of disciplines and academic stages. This diversity was not intended to provide a statistical representation of the graduate student population, but rather to broaden the range of contexts in which academic writing occurs: seminar papers, research proposals, thesis chapters, literature reviews, and manuscripts for submission. Purposive sampling was more suitable here than random sampling, as the study did not seek to ensure an equal probability of inclusion, but rather to verify that participants possessed relevant and sufficiently rich experience for experiential analysis.

Table 1. Participant Characteristics

Task	Context	Demographic Characteristics	Identifier
Writing literature reviews and thesis chapters	Social sciences, research master's	Women, ages 24–31	P01-P02
Course assignments and research proposals	Education and management, master's	Men, ages 27–36	P03-P04
Copyediting, translation, and structuring arguments	Humanities, master's and doctoral	Women, ages 29–42	P05-P06

Task	Context	Demographic Characteristics	Identifier
Drafting academic articles and abstracts	Computer science and engineering, graduate degrees	Men, ages 30–45	P07-P08
Writing empirical papers and discussing findings	Health and psychology, master's	Women, ages 26–38	P09-P10
Developing ideas, feedback, and paraphrasing	Interdisciplinary studies, doctoral	Mixed genders, ages 32–48	P11-P12

4.3. The Interview Guide

The interview guide was designed to balance the uniformity required for cross-participant comparison with an openness to the diverse experiences of using generative AI tools. While it was necessary to maintain common discussion topics, the user experience naturally varies according to discipline, writing confidence, language, and institutional policy. Consequently, a semi-structured protocol was adopted. This protocol covered several fixed themes: how the tools were used throughout the writing process, their perceived contribution, moments of discomfort or apprehension, the sense of textual ownership, the boundaries between assistance and replacement, the understanding of academic integrity policies, and the impact of generative AI on academic writer identity.

The interviews incorporated open-ended and follow-up questions; rather than detailing the full protocol in the body of this paper, the complete guide is provided in the appendix. The questioning strategy focused on eliciting concrete examples: a specific assignment, a particular phrasing generated by the tool, a decision to accept or reject a suggestion, or a dilemma regarding whether to disclose its use. Special attention was paid to questions concerning academic integrity, as these can provoke fears of exposure or judgment. Therefore, it was emphasized at the beginning of the interview that the study's objective was not to detect violations, but rather to understand how students interpret the ethical boundaries within which they operate.

A semi-structured interview format was particularly appropriate for this context, as it allowed for the emergence of unanticipated meanings. Early research on educational chatbots indicates

that enthusiasm for these tools often coexists with concerns regarding unreliability, privacy breaches, manipulation, and academic dishonesty (Tlili et al., 2023). An overly rigid guide might have constrained the conversation to predefined categories, whereas a completely unstructured interview would have hindered systematic comparison across participants. This intermediate structure maintained the research focus without precluding unexpected articulations of the participants' experiences.

4.4. Data Collection

Data collection was designed to allow participants to describe their experiences in detail without feeling evaluated on their compliance with institutional policies. Participants were recruited through targeted outreach in study groups and academic networks, as well as through direct invitations to students who met the inclusion criteria. Before the interview, they received a brief information sheet outlining the study's purpose, the types of questions to expect, the estimated duration, their right to withdraw at any time, and data storage procedures.

Each interview lasted approximately 45 to 60 minutes and took place either via Zoom or in a quiet physical location, depending on the participant's preference. With the participants' consent, the interviews were recorded, transcribed verbatim, and anonymized by altering names, institutions, course titles, and other identifying information. Immediately following each interview, brief field notes were drafted to capture initial impressions, moments of tension, shifts in tone, and recurring themes. Rather than substituting for transcript analysis, these notes served as a supplementary tool to help contextualize and interpret the participants' accounts.

This data collection approach balanced the need for in-depth data with accessibility for busy graduate researchers. In-depth interviews made it possible to explore not only explicit attitudes but also internal contradictions. For instance, a participant might initially describe the tool as being used "strictly for copyediting," yet later reveal how it helped them formulate the central argument of their paper. Such contradictions provided crucial insights for a study focused on ambivalent writing partnerships.

4.5. Data Analysis Method

The analysis sought to balance fidelity to each participant's narrative with the identification of shared patterns across the sample. Therefore, Interpretative Phenomenological Analysis (IPA) was selected, as it is suited for research aiming to understand how individuals make sense of

complex personal experiences. This approach aligned with the research question because using generative AI tools in academic writing is not merely a technical act; it involves relationships among the writer, the text, the institution, academic ethics, and a technology perceived sometimes as a partner and sometimes as a threat.

The analysis began with an in-depth initial reading of each transcript separately, without taking notes, to enter the participant's inner world and avoid premature analysis. Subsequently, exploratory notes were recorded across three dimensions: descriptive, linguistic, and conceptual. At this stage, recurring phrases, metaphors, shifts in phrasing, and points of ambiguity were highlighted—for instance, when participants oscillated between "I wrote" and "the tool wrote for me." From these notes, emergent themes were developed, briefly articulating the psychological meaning of the excerpts.

The next step involved searching for connections within each case. The emergent themes from the first participant were clustered under provisional superordinate headings, such as building writing confidence, fear of losing one's personal voice, and defining boundaries of use. The analysis then moved to the next case, with a conscious effort to bracket the findings from the previous one to avoid imposing preconceived patterns on other participants. Only after completing this idiographic analysis was a cross-case analysis conducted, comparing all cases to identify shared patterns, tensions, differences, and negative cases. Finally, superordinate and subordinate themes were defined and cross-checked against the original transcripts to ensure that every interpretation was grounded in direct quotes.

Interpretative Phenomenological Analysis was more suitable for this study than categorical content analysis because it went beyond merely counting topics to examine the structure of meaning within the experience. This suitability is particularly relevant given the argument that the emergence of large language models disrupts traditional assessment frameworks in higher education, necessitating a rethinking of writing, learning, and evaluation (Rudolph et al., 2023). Consequently, the analysis focused not only on what the students did with the tools but also on how this use related to their self-understanding as academic writers.

4.6. Reflexivity

Reflexivity was essential in this study because the researcher is not an external observer of the phenomenon. As an academic who personally uses generative AI tools for various professional tasks, I approached this research with a positive disposition toward the technology and a curiosity

about its potential. I was aware that this inclination could lead to an overemphasis on experiences of efficiency and empowerment, potentially downplaying participants' feelings of anxiety, dependency, or loss of agency.

To mitigate this bias, I employed bracketing techniques to suspend my preconceptions throughout the interview and analysis stages. Following each interview, I wrote a reflexive memo documenting my emotional reactions, moments of identification with the participant, and instances where I experienced surprise or resistance. Furthermore, during the analysis phase, I deliberately examined excerpts that contradicted the assumption that using these tools primarily empowers the writer. This practice minimized the risk of the analysis becoming a mere confirmation of my preexisting stance. In a broader context, the premise that data technologies are not neutral—but rather entangled with power dynamics and human interpretation—reinforces the need for systematic reflexivity in user experience research as well (D'Ignazio & Klein, 2020).

4.7. Trustworthiness and Rigor

The quality of the research relied on transparency, interpretive consistency, and continuous grounding in the data, rather than on statistical replication. Credibility was strengthened through semi-structured in-depth interviews, the use of direct quotes from the transcripts, and repeated verification of the alignment between the codes and the raw data. Additionally, a peer debriefing session was held for interpretive validation, during which the thematic structure and sample quotes were presented to a colleague specializing in qualitative methods. Whenever possible, member checking was also conducted by sending participants a brief summary of the core interpretations and requesting their feedback.

Transferability was enhanced through a thick description of the participants' characteristics, the institutional context, fields of study, types of assignments, and modes of tool usage. This description does not aim to claim universal generalizability, but rather to enable readers to evaluate whether the findings are applicable to similar contexts. Dependability was reinforced by maintaining a systematic audit trail that documented sampling decisions, the interview guide, interview dates, codebook versions, coding revisions, analytic memos, and the rationale for selecting the final themes. Furthermore, a portion of the coding was re-examined after a period of time to assess interpretive consistency.

Confirmability was bolstered by preserving a clear link between the interview data, the codes, the themes, and the final interpretation. A reflexive journal was used to document

preconceptions, and negative case analysis prevented the erasure of nuances in favor of an overly uniform narrative. These strategies aligned with the interpretive nature of the study: rather than attempting to convert subjective experience into objective measurement, they aimed to elucidate how the interpretation was constructed and why it is well-founded.

4.8. Ethical Considerations

The ethical considerations in this study stemmed from the fact that the topic involves writing practices that institutions might view as sensitive regarding academic integrity. Therefore, the research design had to protect participants from identification, judgment, or the fear that their disclosed information could be used against them. Prior to the interviews, informed consent was obtained, the right to withdraw at any stage was clarified, and it was emphasized that participation was entirely unrelated to any academic evaluation. Participants were asked not to disclose the names of lecturers, courses, or assignments in a way that would enable identification, and any such details were removed from the transcripts if they occurred.

The data were stored in a secure file, with access restricted solely to research purposes. Anonymous identifiers were used in reporting the findings, and quotes were carefully edited to prevent indirect identification without altering their meaning. This ethical handling created conditions for more open discussion of complex, borderline, or ambivalent uses of generative AI tools. Without this protection, participants might have provided only broad and cautious descriptions, and the study would have lost its ability to capture the actual tension between technological assistance, authorial agency, and academic ethics.

5. Findings

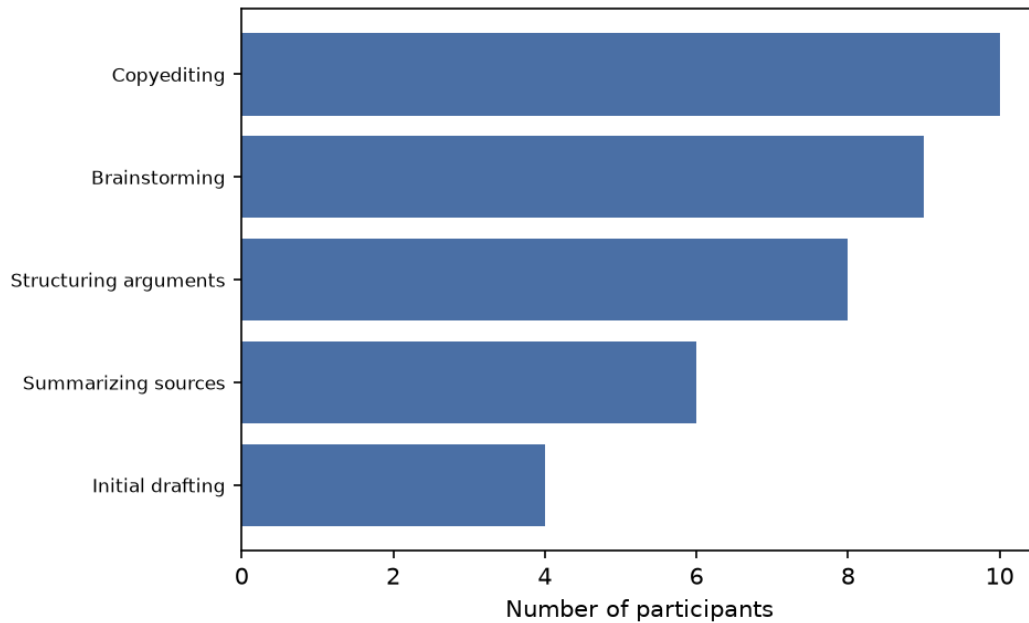
5.1. Participant Profile

The analysis was based on 12 semi-structured in-depth interviews with graduate students who actively used generative AI tools for academic writing tasks. The participants were enrolled in master's or doctoral programs at higher education institutions, representing a disciplinary spectrum that included the social sciences, education, humanities, management, and health sciences. The primary inclusion criterion was not necessarily a high frequency of use, but rather practical experience within the writing process: drafting paragraphs, structuring arguments, copyediting, summarizing sources, preparing research presentations, or receiving feedback on existing text. This selection allowed the focus to remain on how participants interpreted their experiences, rather than merely on their general attitudes toward technology.

Most participants described targeted rather than continuous use: they turned to the tool when encountering writer's block, seeking to improve their academic English phrasing, or needing to quickly organize a complex idea. Only a minority reported the consistent integration of generative AI across all stages of the writing process. This distinction clarifies that the experience did not center on full technological adoption, but rather on an ongoing negotiation with the tool: when it was perceived as a legitimate aid, when as a problematic shortcut, and when as a factor undermining the sense of textual ownership. In this sense, the findings reflect the need for an institutional framework that clarifies responsibility, privacy, training, and usage guidelines, as emphasized by the ecological framework for AI policy in higher education (Chan, 2023).

As shown in Figure 2, the most frequent stages of use were copyediting, brainstorming, and structuring arguments. Fewer participants described using the tool directly to generate a complete draft, and even when such use occurred, it was typically accompanied by critical reading, revision, deletion, or rephrasing.

Figure 2. Primary stages of generative AI use in the writing process



The interviews indicated that using the tool was not perceived as a purely technical action. Even when discussing phrasing corrections or paragraph restructuring, participants linked the action to broader questions of self-efficacy, academic integrity, and researcher identity. Participants asked not only "Did the tool help me?" but also "Is this still my text?", "Am I learning or bypassing the learning process?", and "Do I appear more academic than I really am?" These questions shifted the analysis from the operational level of tool use to the phenomenological level of the user experience, where the tool emerged as an ambivalent writing partner.

5.2. Main Themes

Interpretative phenomenological analysis revealed 4 overarching themes that recurred across the cases, alongside clear differences in their intensity and how participants made sense of them. The first theme, the initial encounter and the tension between curiosity and apprehension, appeared primarily among P1, P2, and P3. It focused on the initial moments of experimentation, where relief over the tool's efficiency coexisted with feelings of guilt and the fear of compromising academic legitimacy. The second theme, shifts in perceived agency and text ownership, emerged among P2, P3, P4, and P5, addressing the question of who leads the writing process when some of the phrasing, structure, or suggestions originate from the tool. The third theme, defining academic identity vis-à-vis the "silent partner," appeared among P1, P2, P4, and P5, exploring how participants redefined expertise, originality, and research capability. The

fourth theme, academic isolation and the need for technological mediation, emerged among P1, P2, P3, and P5, highlighting the tool's role not merely as a writing system, but also as an accessible source of emotional support during an isolating educational journey.

Table 2 presents the main themes, sub-themes, and representative quotes. The table does not replace the interpretative analysis; rather, it consolidates the analytical structure developed through cross-case comparison.

Table 2. Main themes and representative quotes

Representative Quote	Sub-themes	Theme
"There is a constant internal struggle here. On one hand, it's an amazing tool that streamlines processes, and on the other, I ask myself if I am still a real researcher." (P3)	Feelings of guilt and imposter syndrome; amazement at immediate efficiency	The initial encounter and the tension between curiosity and apprehension
"The boundaries are blurred. When half the sentence is mine and half is the tool's, whose text is this ultimately? I feel more like an editor than a writer." (P2)	Loss of personal voice in the writing process; the tool as a catalyst for critical thinking	Shifts in perceived agency and text ownership
"I realized that my expertise is not in knowing how to phrase a sentence in perfect English, but in connecting ideas, in the novelty of my research question." (P4)	Redefining academic expertise; fear of cognitive atrophy	Defining academic identity vis-à-vis the "silent partner"

Representative Quote	Sub-themes	Theme
"My supervisor is very busy and sees me once a month. The AI is available to me at two in the morning, doesn't judge me, and patiently answers every stupid question." (P3)	The computer as an accessible alternative supervisor; longing for personalized human supervision	Academic isolation and the need for technological mediation

The themes are not arranged in a linear sequence. For some participants, the emotional dimension emerged first, followed only later by questions of ownership and integrity; for others, the experience of efficiency paved the way for rethinking their academic identity. However, in all cases involving significant use of the tool, academic writing ceased to be perceived as a solitary, closed activity. It was experienced as a collaborative process, and occasionally as an arena of heightened self-surveillance. This finding echoes the argument that academic writing requires clear guidelines and structured workflows, especially as writing practices evolve and become more complex (Kraus et al., 2022).

5.3. Sub-themes

In the first theme, a sense of guilt stood out, driven primarily by the lack of clear institutional boundaries. P1 described her initial use of the tool: "When I started using it, I felt like I was cheating. It felt as though I wasn't doing the hard work of research, but rather letting a machine write for me." This guilt did not necessarily stem from violating an explicit rule, but rather from the absence of clear institutional guidelines. When participants were unsure whether copyediting, paraphrasing, or initial drafting was permitted, they tended to view the mere act of using the tool as suspicious. For P1, this was compounded by a fear of exposure before an authority figure: "My biggest fear was that my supervisor would find out and think I was lazy, even though I only used it for copyediting." The core experience here was not actual cheating, but rather navigating normative ambiguity.

Alongside this guilt emerged a sense of amazement at the tool's immediate efficiency. P2 noted: "It was simply amazing; suddenly, entire paragraphs I had been stuck on for weeks fell into place in seconds. I felt tremendous relief." This relief was framed not merely as time saved, but as liberation from emotional and cognitive blocks. Participants described instances where a

blank page, cumbersome phrasing, or an unclear structure was rapidly transformed into workable text. In this sense, the tool did not replace the writing process but rather provided a different starting point. This echoes research on artificial intelligence in education, which similarly highlights the value of automated tools for feedback and assessment support, while emphasizing the need for human oversight to ensure accuracy and relevance (Celik et al., 2022).

In the second theme, a shift in the perception of agency and ownership, the dilemma surrounding personal voice was particularly prominent. P4 articulated this sharply: "Sometimes I read what the tool generated and it sounds excellent, but it doesn't sound like me. It feels as though someone else, someone foreign and overly academic, wrote my thesis." This quote highlights a gap between linguistic quality and self-identification. Precisely because the output was "excellent," it elicited discomfort: the excessive polish became a marker of alienation. For other participants, conversely, the tool served to bolster agency. P3 stated: "It gave me back control. Instead of being stuck staring at a blank page, I have a brainstorming partner that enables me to drive the research forward." The distinction between losing and regaining control depended on how the tool was integrated into the writing process: when it drafted on behalf of the participant, a sense of destabilization arose; when it offered options that the participant then selected, rejected, and refined, a sense of ownership emerged.

The sub-theme of the tool as a catalyst for critical thinking further sharpened this point. P5 noted: "Precisely because it sometimes makes mistakes, it forces me to be much more critical. I don't take anything for granted; I argue with the computer." Here, the tool's errors were not perceived merely as failures, but as triggers for verification, skepticism, and conceptual precision. However, not all participants reported a similar capacity. Some admitted that when the generated text sounds convincing, they struggle to distinguish between eloquent phrasing and a valid argument. This exposes the link between AI literacy and academic integrity: it is insufficient merely to know how to operate the tool; one must also be able to evaluate its limitations.

In the third theme, academic identity versus the "silent partner," participants grappled with what remains of their expertise when drafting, translating, and organizing text are partially handled by large language models. P4 described a positive shift: "I realized that my expertise does not lie in knowing how to craft a sentence in perfect English, but in synthesizing ideas and in the novelty of my research question." From this perspective, generative AI does not negate expertise but rather shifts it from the stylistic to the conceptual level. Conversely, P2 expressed a

more profound anxiety: "I worry that if I stop making the effort to articulate things myself, my brain will atrophy. Writing is how I think, and if I surrender that, what else am I surrendering?" Here, writing is not merely a vehicle for presenting pre-existing thoughts, but the very space where thinking crystallizes. The fear of cognitive atrophy thus reflects an anxiety over the loss of a process, not just a product.

Academic identity was also tied to arenas where students were required to stand behind their written work. P1 linked this to public exposure: "When I present at a conference, I suddenly feel exposed. Is the audience evaluating me, or the enhanced version that the AI generated for me?" The experience does not end with the submission of an assignment; it extends into evaluative spaces where students must embody and voice their texts. In contrast, P5 offered a broader perspective: "It forces me to redefine what it means to be a researcher in the 21st century. We can no longer afford to be left behind." This stance views the use of such tools as part of a broader professional evolution rather than a deviation from academic norms. At the same time, caution is warranted against drawing overly sweeping technological conclusions: while the technical literature on generative models describes the growing use of techniques to synthesize information or examples when original data is scarce, it does not directly address the writing experiences of graduate students (Alzubaidi et al., 2023).

In the fourth theme, academic loneliness and the need for technological mediation, the tool emerged as an accessible presence within an emotional landscape characterized by overload, doubt, and a lack of guidance. P5 noted: "Writing a thesis is an incredibly lonely process. Suddenly, with the chat open, it feels as though someone is there with me in the dark study." P2 added: "It has become a sort of academic psychologist for me. I vent my research frustrations there, and it helps me reorganize my thoughts." These quotes point to a clear departure from the narrow definition of a writing tool. Participants sought more than mere copyediting; they sought responsiveness, presence, patience, and the temporary validation needed to keep going.

Yet, this availability did not eliminate the yearning for human mentorship. P1 remarked: "With all due respect to artificial intelligence, it can't give me a pat on the back or tell me it's proud of me. I deeply miss that human warmth." The distinction between automated responses and human recognition was central: while the tool can reply, it cannot assume pedagogical responsibility or provide an interpersonal relationship. Consequently, the user experience

described in the interviews was twofold: on one hand, it alleviated loneliness; on the other, it starkly highlighted the deficiencies in the human supervisory framework.

5.4. Outlier Cases

Alongside the main patterns, one outlier case emerged, primarily in the accounts of P5, challenging the interpretation of an agency crisis or a loss of personal voice. As they noted: "I don't understand what loss of control everyone is talking about. For me, it is exactly like using a calculator or autocorrect. The text is 100% mine, and the tool is just an advanced pen." This perspective does not reject the use of the tool itself, but rather the psychological baggage that other participants attached to it. From this viewpoint, generative AI is a technical extension of an existing capability, not a co-author or an identity-shaping agent.

This outlier prevents an overly uniform portrayal of student perceptions. It demonstrates that the same action—such as rephrasing a paragraph—might be experienced by one participant as a compromise of authenticity, and by another as a neutral, technical enhancement. This difference stems not from the tool itself, but from how the student interprets academic writing: if writing is viewed as a process of personal conceptualization, the tool's intervention may feel destabilizing; however, if writing is seen as the efficient production of text to express an idea the author already owns, the tool is regarded simply as a piece of equipment.

In the cross-case analysis, the outlier did not invalidate the 4 themes but rather refined them. It revealed that ambivalence is not an inevitable outcome of every use; rather, it is a prominent feature when the tool enters areas where students identify as writers, researchers, and possessors of an academic voice. Therefore, the central finding is not that all participants experience a loss of agency, but rather that the use of generative AI exposes the conditions under which agency, ownership, and academic integrity become active concerns during the writing process.

6. Discussion

6.1. Interpretation of the Main Findings

The findings indicate that graduate students neither relegate generative AI tools to the periphery of the writing process as mere technical aids, nor do they entirely outsource the task of writing to them. The interviews revealed an experience of a tense partnership: the tool is perceived as an accessible source for ideas, alternative phrasing, initial argument structuring, and copyediting, while simultaneously acting as an entity that requires human verification, restraint, and decision-making. Consequently, its contribution extends beyond mere time-saving. It affects how graduate students perceive themselves as writers, how they define the boundaries of legitimate assistance, and how they view their responsibility for the text they submit.

In this sense, the use of generative AI shifts the focus of academic work to some extent. Rather than being perceived as a linear sequence of planning, drafting, and editing, the writing process emerges as a space for dialogue, revision, and oversight. Students do not merely "receive" text from the tool; they scrutinize it, reject certain parts, adopt specific phrasing, and renegotiate their relationship with the argument. These actions indicate that academic agency does not necessarily vanish with the introduction of the tool, but rather transforms: it shifts from the ability to generate text from scratch to the capacity to evaluate, filter, justify, and take responsibility for the final product. However, this shift is not neutral. When the tool offers fluent, authoritative, and sometimes overly convincing phrasing, it risks blurring the distinction between comprehension and linguistic production, as well as between stylistic improvement and the construction of a research argument.

This ambivalence was particularly pronounced regarding academic integrity. On the one hand, students described using the tool as an extension of existing practices, such as consulting with peers, using a dictionary, copyediting, or receiving feedback. On the other hand, they noted that generative AI operates on a different scale: it can generate argumentative sequences, entire paragraphs, and phrasing that resembles cohesive academic writing. This creates a difficulty in defining the threshold between legitimate assistance and actual participation in the writing process. This difficulty is not merely normative but also experiential. Students may feel that the text is "theirs" in terms of responsibility and intent, yet not entirely theirs regarding the specific

wording, structure, or initial conceptual direction. This non-binary sense of ownership is one of the study's most prominent findings.

A potential counterargument is that emphasizing risk, ambivalence, and integrity might overshadow the contribution of generative AI to graduate students, particularly in contexts of heavy research workloads, writing in a non-native language, or struggling to initiate a complex task. Indeed, the findings do not justify an attitude of rejection or anxiety toward the tool. On the contrary, they demonstrate that for some students, the tool lowered barriers to academic writing, enabled rapid conceptual experimentation, and provided immediate feedback when human feedback was unavailable. However, these practical benefits do not negate the questions of responsibility. Frequent use or a positive attitude toward the tool does not necessarily indicate its critical, ethical, and stable integration. Therefore, the findings are neither "for" nor "against" generative AI; rather, they illustrate that the significance of its use is determined by the users' quality of judgment, their level of AI literacy, and the institutional clarity surrounding the writing process.

6.2. Comparison with the Literature

The findings align with literature describing large language models as highly capable of producing human-like text, yet also as technologies that raise unresolved questions in academic contexts. Roumeliotis and Tselikas's review demonstrates that these models rely on natural language processing, supervised learning, and reinforcement learning, enabling them to generate text that closely resembles human output; however, this very capability also creates challenges regarding trust, verification, and responsible use (Roumeliotis & Tselikas, 2023). In the context of the current study, this explains why students experience the tool as particularly effective during the drafting and editing stages, but also as destabilizing precisely because the text looks "good" before its accuracy, originality, or alignment with the research argument have been verified. In other words, rhetorical fluency is no guarantee of epistemic quality.

This comparison also highlights the tension in the literature between viewing generative AI as a promoter of thinking and the concern that it weakens critical thinking through overreliance. The current findings do not resolve this tension by measuring writing quality or critical thinking, but they do allow it to be interpreted from an experiential perspective. When students used the tool to open up possibilities, test formulations, identify weaknesses, or receive initial feedback, it likely expanded their cognitive space. Conversely, when they adopted ready-made output simply

because it sounded academic, a portion of the cognitive work was shifted from the writer to the system. Therefore, the question is not whether the tool "develops" or "weakens" thinking, but rather under what conditions it is integrated into the writing process. The finding supports a conditional conclusion: generative AI can contribute to writing when it operates within a practice of checking, redrafting, and justifying choices, but it may encourage dependency when it replaces the work of decision-making.

The literature on AI in higher education also points to the importance of institutional and ethical conditions. Bond and colleagues note that research in this field often focuses on adaptive systems and personalization, yet lacks sufficient engagement with ethical, methodological, and institutional contexts (Bond et al., 2024). This finding is particularly relevant because the students in the current study described their user experience not only in terms of convenience or efficiency but also in terms of ambiguity: What is permitted? What requires disclosure? Is language editing with the tool equivalent to receiving human assistance? Does an idea suggested by the tool belong to the writer if it has undergone significant revision? These questions cannot be resolved solely at the individual level. When policies are unclear, the responsibility shifts to the students, but without providing them adequate tools to distinguish between legitimate and problematic use.

Chiu's study, although conducted among teachers and educational leaders rather than graduate students, also helps frame the findings. It highlights the need for a new type of foundational knowledge, interdisciplinary instruction, and a reassessment of evaluation methods and professional standards following the introduction of generative AI tools into education (Chiu, 2023). The parallel to the current study is clear: students need more than just technical guidance on how to write an effective prompt; they require a broader understanding of the interrelations among technology, text, source, responsibility, and discipline. AI literacy, in this context, is not merely an operational skill; it also encompasses the ability to recognize output limitations, assess suitability for academic genres, distinguish between paraphrasing and altering an argument, and understand when using the tool requires disclosure.

However, caution is needed to avoid overgeneralizing the findings. The current study is based on 12 in-depth interviews; therefore, its strength lies not in presenting usage rates or making statistical determinations between groups, but in identifying patterns of meaning. Furthermore, interview and self-report data may reflect self-presentation, fear of institutional judgment, or

retrospective rationalization. This limitation does not negate the value of the findings, but it clarifies their status: they reveal how students understand, explain, and justify their use of generative AI tools, rather than providing a complete observation of all actual writing practices. Precisely for this reason, there is significance in the gaps that emerged in their accounts between recognizing the utility of these tools and fearing the compromise of ownership, trust, and academic integrity.

6.3. Research Contribution and Practical Implications

The contribution of this study lies in shifting the discussion on generative AI tools from the technical question of whether they are a problem or a solution, to examining them as a component that shapes the academic writing experience. While public and institutional discourse often focuses on whether students are "cheating" or "becoming more efficient," this research demonstrates that the experience is more complex: graduate students navigate a space where assistance, dependence, responsibility, ownership, and apprehension coexist. Consequently, the study contributes to understanding student perceptions not as broad pro- or anti-technology stances, but as interpretations constructed around concrete writing events: drafting a paragraph, rewriting a claim, checking for clarity, seeking structure, or deliberating whether to disclose the use of a tool.

Theoretically, the study reinforces the argument that academic writing in the era of generative AI cannot be understood solely through concepts of originality or efficiency. It must also be examined through the lenses of agency, authority, and trust. The student is not merely a producer of text, but rather manages a relationship with a non-human textual source that sometimes behaves as a conversational partner. These relationships destabilize familiar categories of assistance, copyediting, guidance, and copying. Therefore, the relevant distinction is not between "human" and "artificial" writing, but between varying levels of involvement, control, comprehension, and responsibility. This contribution integrates into the broader literature on writing studies, research methods, and academic ethics, demonstrating that questions of phrasing are inseparable from questions of methodology, and that the use of writing tools is not merely a stylistic matter but part of the conditions of academic knowledge production.

The practical implications directly concern higher education institutions, supervisors, lecturers, and writing support centers. Institutional policies must be clearer and move beyond general prohibitions. A blanket ban may drive usage underground and exacerbate distrust

between students and the institution. Conversely, unrestricted permission without clear distinctions risks blurring the boundaries of academic integrity. What is required, therefore, is a tiered classification of uses: language assistance, brainstorming, structuring, copyediting for clarity, drafting, integrating sources, and formulating an argument are not ethically or pedagogically equivalent. Such distinctions enable students to understand not only what is permitted, but why.

Academic writing instruction should integrate generative AI literacy as a core component of research training. The goal is not to turn every writing course into a technology class, but rather to teach students how to critically read generative AI output: verifying sources, identifying empty generalizations, assessing alignment with the theoretical framework, distinguishing between improved phrasing and altered arguments, and drafting appropriate disclosures when required. For graduate students, these skills are not an external add-on, but an integral part of their developing academic professionalism as researchers and writers.

Supervisors also play a crucial role in fostering open dialogue about tool usage. Instead of assuming that the use of generative AI inherently signals weakness or a lack of integrity, supervisors can ask students which stages of the process involved the tool, what decisions were made based on its output, and what aspects remain under the student's direct responsibility. Such questions cultivate reflexive responsibility rather than settling for mere surveillance. They also help identify instances where the tool is particularly beneficial: overcoming writer's block, organizing extensive material, obtaining initial feedback, or improving clarity. At the same time, they establish a clear boundary when the tool replaces reading, thinking, source analysis, or argument construction.

Finally, the study suggests viewing generative AI neither as an external threat to academic writing nor as a magic bullet for its challenges, but as an arena that re-exposes the fundamental principles of writing: responsibility for claims, transparency regarding the process, understanding of sources, and the ability to explain why the text is written the way it is. In doing so, it shifts the center of gravity back from the tool to the writer. A partnership with the tool can be beneficial, but only as long as the student remains the one who understands, chooses, verifies, and bears academic responsibility for the final product.

7. Limitations of the Study

The primary limitation of the research design stems from its qualitative, observational, and retrospective nature. The interviews documented how graduate students retrospectively reconstructed and interpreted their use of generative AI tools, rather than capturing real-time usage. Consequently, the findings are particularly suited for identifying meanings, justifications, and interpretive patterns, but less so for determining actual behavior. They may therefore overestimate the degree of order, awareness, and control participants attributed to themselves during the writing process, while underestimating moments of experimentation, confusion, concealment, or uses that conflict with academic integrity. However, this retrospective approach is not merely a weakness; it allowed participants to present their use of the tools as part of the broader writing process. This advantage does not negate the limitation regarding causality. The study was not designed as a systematic multi-institutional comparison and did not include a control group or experimental conditions. In a qualitative design of this kind, blinding and experimental control do not apply in the conventional sense; therefore, the findings cannot be used to infer a causal effect of the tools on writing quality or academic achievement.

Regarding the population and sample, the study was based on 12 in-depth interviews with graduate students who actively used generative AI tools in their academic writing. While the sample size aligns with the qualitative logic of interpretive depth, it limits generalizability: the findings highlight plausible possibilities and patterns rather than their prevalence in the wider population. Another notable limitation is self-selection bias. Those who agreed to be interviewed on this topic may be students with heightened awareness, a willingness to discuss academic ethics, or relative confidence in their ability to explain their use of the tools. The likely direction of this bias is thus an overestimation of reflexivity and controlled use, alongside an underestimation of covert, dependent, or potentially embarrassing uses. Conversely, highly aware participants might actually be more critical in describing their dilemmas, thereby emphasizing the risks. This possibility is also plausible, but it does not alter the methodological conclusion: the sample is suited for understanding mechanisms of meaning-making, not for estimating the scope of the phenomenon. Attrition is not a significant limitation in this study, as participation involved a single interview session rather than longitudinal follow-up.

In terms of measurement and data, the study relied on self-reporting and did not include usage logs, version histories, writing samples, or system data. Consequently, it was not possible to

compare participants' accounts with their actual practices—for instance, whether a specific tool was used for targeted copyediting or for broader argument formation. This gap lies at the heart of the research problem, as the distinction between "language assistance" and "conceptual contribution" is not merely a technical measurement detail, but rather integral to how participants understand their use of the tools. This limitation may lead to an underestimation of the depth of the tools' involvement in the text, particularly when participants sought to present themselves as writers with full agency. Furthermore, AI literacy was discussed here as an interpretive condition emerging from the interviews, rather than being measured as a mediating variable using a validated instrument. To mitigate this limitation, follow-up questions were used to elicit concrete examples of different stages in the writing process, rather than relying solely on general attitudes.

Regarding the analysis, the interpretation was based on the thematic coding of verbal testimonies and is therefore sensitive to the researcher's interpretive decisions. Any subgroup comparisons that emerged during the analysis were exploratory rather than statistical; thus, issues of multiple comparisons do not apply here in the quantitative sense. Nevertheless, some blurring remains among academic discipline, institutional policy, level of experience in academic writing, and the experience of using different tools. The direction of the bias is ambiguous: these factors might amplify feelings of ambivalence for some participants while obscuring them for others. Therefore, the conclusions were carefully formulated as claims about how participants understood and managed their use of the tools, rather than as sweeping generalizations about the entire student population or the efficacy of the tools themselves. To moderate this limitation, a clear distinction was maintained between describing participants' experiences and drawing broad conclusions about the wider student population.

8. Recommendations for Future Research

Future research should build on the insight that emerged in this article: the use of generative AI tools in academic writing is not merely a matter of efficiency, but also involves the reconfiguration of responsibility, ownership, trust, and agency. Since the current study was based on 12 semi-structured in-depth interviews, this investigation should be expanded using multi-site qualitative research designs encompassing various institutions, diverse disciplines, and different stages of study among graduate students. Such an expansion is intended not only to increase the sample size but also to test the generalizability of the findings. If the user experience is similar among humanities thesis writers, social science research-track students, graduate researchers in

the exact sciences, and students in professional tracks, it will be possible to describe a broad pattern of shifting academic literacy; if not, it will be necessary to explore how each disciplinary context generates distinct norms of assistance, suspicion, and legitimation.

Another direction for research is a shift from retrospective to longitudinal studies that follow students throughout the writing process itself. Post hoc interviews yield rich interpretive insights, but they rely on memory, self-justification, and the ways participants wish to present themselves as responsible researchers. Therefore, future studies could incorporate writing diaries, documenting text revisions, feedback sessions with supervisors, and analyses of decision points where students choose to use or refrain from using generative AI tools. Such an approach would make it possible to distinguish between normative statements and actual practices, examining not only what students say about their usage but also how this usage actually unfolds across the sequence of planning, drafting, copyediting, and reviewing.

There is also room to examine AI literacy as a developing variable rather than a fixed trait. While the current study noted that students distinguish between usage perceived as helpful and usage seen as compromising academic integrity, the mechanisms through which these distinctions are acquired remain unexplored. Future studies should therefore systematically examine how training, institutional policy, instructor guidance, and accumulated experience relate to students' ability to evaluate reliability, identify biases, establish usage boundaries, and declare the tool's involvement in the text. One might argue that the responsibility for appropriate use is primarily personal; however, the study's findings indicate that moral and practical distinctions are shaped within an institutional environment, making it necessary to also examine the learning conditions that produce them.

A complementary direction is to broaden the research focus beyond students to the academic system in which they operate. Studies incorporating supervisors, lecturers, ethics committees, and policymakers could explore how implicit expectations surrounding the use of generative AI tools are formed, and how gaps between official policy and daily practice relate to students' user experiences and perceptions. This direction is essential for developing policies that go beyond blanket prohibitions or permissions, instead defining situations, responsibilities, and transparency requirements according to the nature of the assignment and the stage of the writing process. In this way, future research can move from describing user attitudes to understanding the

institutional mechanisms that either facilitate responsible use or, conversely, perpetuate ambiguity and insecurity.

9. Summary and Conclusions

9.1. Summary of Main Findings

This study examined a research problem brought into sharper focus by the rapid integration of generative AI tools into daily learning and research routines: how should we understand the role of these tools in graduate students' writing processes, given that they are not merely technical aids but are embedded in the space where thinking, responsibility, textual ownership, and academic integrity take shape? The argument developed throughout this work is that graduate students do not view generative AI tools as a simple substitute for the academic writer. Rather, they position them as an ambivalent writing partner: a source of assistance, acceleration, drafting, and feedback, but also a source of anxiety regarding dependency, blurred authorship boundaries, and diminished trust in the text.

The analysis revealed that these tools are used primarily within the writing process itself, rather than outside it. Interviewees described using generative AI to refine ideas, outline arguments, rephrase paragraphs, perform copyediting, and check text clarity. In these respects, the tool is perceived as augmenting the student's capabilities, not just as a shortcut. However, the same action took on different meanings depending on the degree of control the student felt over the text. When the tool's suggestion served as a starting point for critical examination, it was experienced as legitimate assistance; when it replaced judgment or created a sense that the text was "already written," it provoked discomfort regarding ownership and academic integrity.

Another finding relates to the gap between efficiency and trust. Participants recognized the practical value of the tools, especially when facing heavy workloads, drafting difficulties, writing in a non-native language, or needing rapid feedback. At the same time, they noted that the outputs should not be treated as an independent source of authority. Relying on large language models required them to verify, filter, compare, and rephrase. Consequently, using generative AI tools did not eliminate the writer's responsibility but shifted it to a more complex level: the responsibility not only to write but also to understand how the text was produced, what assumptions are embedded within it, and what the relationship is between the tool's contribution and the student's own.

The study also highlighted the role of the institutional context. The students did not operate in a normative vacuum; they interpreted their use of the tools through the lens of instructors' expectations, ambiguous guidelines, the nature of the assignment, and the ethical climate of their academic environment. This finding aligns with studies showing that students tend to view generative AI as a means to increase productivity and personalize learning, whereas faculty emphasize pedagogical and ethical concerns alongside the need for clear institutional policies (Chan & Lee, 2023). Therefore, the student experience is not merely personal or technological; it is shaped by the relationships of trust between students, faculty, and the institution, as well as by standards of academic ethics.

9.2. Conclusions

The first conclusion to emerge from this study is that the use of generative AI in academic writing cannot be understood through a binary division between "permitted use" and "cheating." This dichotomy fails to capture the reality revealed in the research. Students do not describe these tools as inherently usurping their role as writers, nor do they view them as entirely neutral instruments. The academic significance of their use is determined by the practice itself: the stage of the writing process at which the tool was employed, what the student sought to achieve, how the output was evaluated, and whether the writer maintained intellectual responsibility for the final text.

A second conclusion follows: AI literacy is a prerequisite for responsible use, but it should not be reduced to the technical skill of prompting. This literacy encompasses an understanding of the limitations of large language models, the ability to evaluate their reliability, the capacity to distinguish between stylistic assistance and conceptual contribution, an awareness of potential biases, and the ability to articulate the tool's role in the writing process. In other words, students must develop a reflexive stance toward its use: not merely knowing how to "operate" the tool, but knowing when to reject its suggestions, when to rewrite them, and when to avoid them entirely because they weaken the argument or obscure their research voice.

The third conclusion concerns the concept of agency. The study indicates that student agency is not measured solely by whether they wrote the sentences themselves, but by their ability to lead the writing process, make decisions, justify choices, and take responsibility for the final product. Generative AI tools can enhance agency when they enable students to overcome drafting hurdles, explore alternatives, or receive preliminary feedback. However, they can diminish it

when students adopt the output without sufficient understanding, or when the desire to streamline the work replaces conceptual processing. Therefore, agency is not a fixed attribute but a pedagogical achievement that must be cultivated.

The fourth conclusion is that the discourse on academic integrity must shift from a logic of mere detection and deterrence to one of guidance, transparency, and accountability. This does not mean abandoning boundaries; on the contrary, it requires establishing clearer boundaries that distinguish between different uses and clarify what is expected of students when integrating the tool into their work. In this context, the literature emphasizes that the responsible integration of tools such as ChatGPT requires institutional leadership, ethical frameworks, and authentic assessment, rather than settling for an approach focused solely on detecting anomalies (Crawford et al., 2023). This conclusion aligns with the study's findings: ambiguity does not create academic freedom but rather insecurity, and sometimes even encourages covert use instead of open, informed discussion.

A limitation of this study lies in its deliberate focus on graduate students in higher education institutions who actively use generative AI tools in academic writing assignments. This choice allowed for an in-depth examination of the experiences of writers at a stage where writing is not merely a coursework assignment, but also part of their development as researchers. Consequently, the study's conclusions are primarily valid for understanding the user experience and the meaning students attribute to using these tools in this specific context, rather than being generalizable to all student populations or all types of academic assignments.

9.3. Recommendations

The conclusions indicate a need to formulate an institutional policy that goes beyond general prohibitions or declarations of principle. Higher education institutions must offer graduate students detailed guidelines clarifying which uses of generative AI are permitted, which require disclosure, and which cross the boundaries of academic integrity. These guidelines should be tailored to the nature of the assignments: using a tool for copyediting is not the same as generating a theoretical argument, and receiving general feedback is not the same as producing an entire section of a paper. An effective policy will recognize this complexity and allow students to act with confidence, without treating every use of the tool as inherently suspicious.

Alongside policy, it is recommended to integrate systematic training in AI literacy as an integral part of graduate training. Such training should not be viewed merely as a technical

workshop, but as a component of academic writing instruction. It should address evaluating outputs, drafting usage declarations, identifying errors, and maintaining an independent research voice even when relying on an external tool. In this way, students can be supported in transforming the use of generative AI from an intuitive and sometimes covert practice into a conscious, controlled, and justifiable one.

A further recommendation concerns assignment design and assessment. If institutions wish to preserve the value of academic writing, they must design assignments that emphasize process, reflection, and reasoning, rather than just the final product. Students could be asked to explain how they developed their argument, which sources shaped their position, at which stages they utilized generative AI tools, and what revisions they made following this use. Such requirements are not intended to be burdensome, but rather to recenter the author's responsibility. When students are required to account for their writing process, the very use of the tool does not negate learning, but may instead become a part of it.

Furthermore, it is recommended that discussions regarding generative AI tools take place at the level of the academic unit and discipline, rather than solely at the level of general institutional regulations. Different fields of study assign varying degrees of importance to style, interpretation, data analysis, conceptual originality, and disciplinary writing. Therefore, an overly uniform policy may overlook the distinctions between different types of writing. A broad institutional framework should establish binding principles while allowing each discipline to translate them into practical guidelines that align with its specific characteristics of knowledge, assessment, and ethics.

Ultimately, the contribution of this study lies in presenting a nuanced picture of student perceptions regarding generative AI in academic writing. The question is not whether these tools are "good" or "dangerous," but rather under what conditions they enhance learning, writing, and agency, and under what conditions they undermine them. This work marks a transition from a generalized observation of the technology to a research-based understanding of the user experience within an actual academic context. The professional and research trajectory that emerges from this is the development of training and assessment frameworks that will assist graduate students in using generative AI tools transparently, critically, and responsibly, without compromising the fundamentals of academic writing or their accountability as emerging writers and researchers.

References

- Alzubaidi, L., Bai, J., Al-Sabaawi, A., Santamaría, J., Albahri, A. S., Al-dabbagh, B. S. N., Fadhel, M. A., Manoufali, M., Zhang, J., Al-Timemy, A. H., Duan, Y., Abdullah, A., Farhan, L., Lu, Y., Gupta, A., Albu, F., Abbosh, A., & Gu, Y. (2023). A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications. *Journal of Big Data*, 10(1), Article 46. <https://doi.org/10.1186/s40537-023-00727-2>
- Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, 21(1), Article 4. <https://doi.org/10.1186/s41239-023-00436-z>
- Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., Bamber, G. J., Beltran, J. R., Boselie, P., Lee Cooke, F., Decker, S., DeNisi, A., Dey, P. K., Guest, D., Knoblich, A. J., Malik, A., Paauwe, J., Papagiannidis, S., Patel, C., Pereira, V., Ren, S., ... Varma, A. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. *Human Resource Management Journal*, 33(3), 606–659. <https://doi.org/10.1111/1748-8583.12524>
- Carvalho, D. V., Pereira, E. M., & Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. *Electronics*, 8(8), 832. <https://doi.org/10.3390/electronics8080832>
- Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends*, 66(4), 616–630. <https://doi.org/10.1007/s11528-022-00715-y>
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), Article 38. <https://doi.org/10.1186/s41239-023-00408-3>
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), Article 43. <https://doi.org/10.1186/s41239-023-00411-8>

- Chan, C. K. Y., & Lee, K. K. W. (2023). The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers? *Smart Learning Environments*, 10(1), Article 60. <https://doi.org/10.1186/s40561-023-00269-3>
- Chiu, T. K. F. (2023). The impact of Generative AI (GenAI) on practices, policies and research direction in education: a case of ChatGPT and Midjourney. *Interactive Learning Environments*, 32(10), 6187–6203. <https://doi.org/10.1080/10494820.2023.2253861>
- Crawford, J., Cowling, M., & Allen, K.-A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching and Learning Practice*, 20(3). <https://doi.org/10.53761/1.20.3.02>
- Dergaa, I., Chamari, K., Zmijewski, P., & Ben Saad, H. (2023). From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, 40(2), 615–622. <https://doi.org/10.5114/biolsport.2023.125623>
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2019). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>
- Hanelt, A., Bohnsack, R., Marz, D., & Antunes Marante, C. (2020). A systematic review of the literature on digital transformation: Insights and implications for strategy and organizational change. *Journal of Management Studies*, 58(5), 1159–1197. <https://doi.org/10.1111/joms.12639>
- Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2023). Interpreting black-box models: A review on explainable artificial intelligence. *Cognitive Computation*, 16(1), 45–74. <https://doi.org/10.1007/s12559-023-10179-8>
- Kraus, S., Breier, M., Lim, W. M., Dabić, M., Kumar, S., Kanbach, D., Mukherjee, D., Corvello, V., Piñeiro-Chousa, J., Liguori, E., Palacios-Marqués, D., Schiavone, F., Ferraris, A., Fernandes, C., & Ferreira, J. J. (2022). Literature reviews as independent studies:

- guidelines for academic practice. *Review of Managerial Science*, 16(8), 2577–2595.
<https://doi.org/10.1007/s11846-022-00588-8>
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and Opportunities of Generative AI for Higher Education as Explained by ChatGPT. *Education Sciences*, 13(9), 856.
<https://doi.org/10.3390/educsci13090856>
- Park, S.-M., & Kim, Y.-G. (2022). A metaverse: Taxonomy, components, applications, and open challenges. *IEEE Access*, 10, 4209–4251. <https://doi.org/10.1109/access.2021.3140175>
- Roumeliotis, K. I., & Tselikas, N. D. (2023). ChatGPT and Open-AI models: A preliminary review. *Future Internet*, 15(6), 192. <https://doi.org/10.3390/fi15060192>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, 6(1).
<https://doi.org/10.37074/jalt.2023.6.1.9>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), Article 15.
<https://doi.org/10.1186/s40561-023-00237-x>

Appendices

Appendix A: Interview Guide

This guide serves as a **draft / illustrative tool for a semi-structured in-depth interview**, formulated based on the study's interview protocol. Because no actual data have been collected yet, it should be considered a proposed instrument that must be piloted with at least one participant prior to data collection. This is to ensure that the questions are clear, not overly leading, appropriate for the interview duration, and enable participants to describe in depth their experiences using generative AI tools in academic writing. Estimated interview duration: approximately 60 minutes.

Informed Consent and Ethical Preamble

The following consent statement must be read to the participant either before or at the start of the recording, in accordance with ethical approval protocols:

"Hello, and thank you for agreeing to participate in my research, which explores the experiences of graduate students using generative AI tools during the academic writing process. This interview will be recorded solely for transcription and analysis. Your participation is entirely voluntary; you may end the interview at any time or decline to answer specific questions. Your identity will remain strictly anonymous. Any identifying details will be removed or altered in the final research report, and the recordings will be stored in a secure folder and deleted upon completion of the study. Do you consent to participate in this research and to have the interview recorded under these conditions?"

Interviewer Guidelines for Sensitive Topics

Because the interview addresses academic integrity, text ownership, and the use of artificial intelligence tools in academic assignments, interviewers must emphasize to participants that the interview is not intended to evaluate, judge, or detect violations of institutional policies. Rather, the objective is to understand their experiences, dilemmas, and the meanings they attribute to using these tools. If a participant expresses discomfort, they should be permitted to skip the question, provide a more general response, or take a short break.

Introductory and Background Questions

1. Briefly tell me about your graduate research field and what a typical writing day has looked like for you recently.
2. When and how were you first introduced to generative AI tools in the context of your academic work?

Core Questions

1. Describe a specific instance where you used a generative AI tool while working on an academic text you recently wrote.
 - At which stages of the writing process—for example, brainstorming, drafting, or copyediting—did you choose to integrate the tool in that instance?
 - What were the instructions, requests, or prompts you provided to the tool at that stage?
2. What reasons or needs led you to use these tools in your writing processes?
 - What writing difficulties or blocks did you hope the tool would help you overcome?
 - How have your motivations for using the tool changed over time as you gained more experience with it?
3. How do you experience the dialogue or interaction with the interface while working on an academic text?
 - How did you respond in situations where the output you received from the tool was inaccurate or did not align with your academic style?
 - What process do you go through to make the generated output an integral part of your writing?
4. What meaning or value do you attribute to a text written with the assistance of AI compared to a text you wrote without such assistance?
 - How do you currently perceive the concept of "originality" in the texts you submit?
 - What feelings or thoughts arise regarding your intellectual ownership of the final text?
5. How, if at all, has the integration of generative AI tools affected your sense of self-efficacy and confidence as an academic writer?
 - To what extent do you feel that control and agency over the writing process remain in your hands?

- How has your assessment of your personal capabilities as a researcher changed following the use of these tools?

6. How does the use of these tools reshape the way you perceive your academic identity as a graduate student?

- What boundaries do you set for yourself when using AI to still feel like a "real researcher"?

- How do you think the academic community around you perceives students who use these tools, and how does this affect you?

Concluding Questions

1. Is there any other aspect of your experience writing with AI that I have not asked about, which you would like to share or elaborate on?

2. If I need a brief clarification during data transcription and analysis, may I contact you again via email?