AJELS can be found on EBSCOhost databases
https://doi.org/10.59960/12.2.a5

Design, Methods, and Theories in Research on Writing Competence Development in the GenAI Writing Era

Nikko J. Domingo & Rachelle B. Lintao

Received: September 14, 2024
Reviewed: October 21, 2024
Revised: November 3, 2024
Accepted: November 19, 2024

Corresponding author
Nikko J. Domingo
University of Santo Tomas, the Philippines
domingonikko985@gmail.com

Rachelle B. Lintao
University of Santo Tomas, the Philippines

Additional information is available at the end of the article

Abstract

This study presents a systematic review of the designs, methods, and theories used in empirical studies on AI-powered digital writing tools (AI-DWTs) within the context of Technology-Mediated Language Learning (TMLL). Despite the numerous systematic reviews on AI in education (AIED), reviews focusing on the designs, methods, and theoretical approaches used in investigating AI-DWTs in developing writing competence remain underexplored. Hence, this study addresses three research questions: (a) What research designs are used in studies on AI-DWTs for teaching and learning writing? (b) What research methodologies, including data collection and mode of analysis, are used in studies on AI-DWTs for teaching and learning writing? and (c) What theories/models are most commonly employed in research on AI-DWTs for teaching and learning writing? The study found several experimental designs, mostly utilizing quantitative and mixed methods. Pre-tests and post-tests through writing tasks and questionnaires for assessing perceptions were prevalent data collection techniques. Thematic analysis was employed for data from survey and systematic review designs, while paired sample t-tests and independent-sample t-tests were used for analyzing test scores. However, most studies that implicitly leaned toward TMLL often lacked a solid theoretical foundation. This perennial challenge has implications for future research on AI-DWTs if not addressed. For future studies on AI-DWTs in enhancing writing competence, it is recommended to focus on direct observation of learner-AI-DWT interactions (LAII) as well. This would examine how learners engage with these tools, prompt them, employ strategies, process information, and how these interactions influence the recursive stages of the writing process.

Keywords: methodological/theoretical review, AI-DWTs, TMLL, teaching/learning writing competence

https://doi.org/10.59960/12.2.a5

1. Introduction

In language education, writing is one of the fundamental competencies that teachers must develop among students. This competence includes a complex process that combines basic skills such as spelling and mechanics with higher-level abilities like content organization, logical sequencing, and stylistic appropriateness (Godwin-Jones, 2022). For L2 learners, this competence is even more challenging due to potential gaps in vocabulary, grammar, pragmatic understanding, or rhetorical skills. As a result, many L2 learners use digital and online resources to support their writing efforts. Eslit (2023) elaborates on how technology has been seamlessly integrated into English language teaching and learning (ELT/L) through various platforms, which enhance student engagement and improve learning outcomes. Stapleton and Radia (2010) further describe this as the Tech-era L2 writing characterized by the introduction of tools like word processors and online resources in the late 1970s (Pennington, 1993). In recent years, technology's role in writing has focused on *automated writing evaluation* (AWE). Nowadays, with the deployment of AI-powered Digital Writing Tools (henceforth, AI-DWTs) powered by Generative AI (GenAI), which can generate entire texts from brief prompts, advances in deep learning and larger datasets have further transformed the field of teaching and learning writing.

The role of GenAI-powered writing tools in developing writing skills is becoming increasingly prominent (Crompton et al., 2022). Hence, this area of research has been explored in various studies, which have revealed its influence on multiple aspects of writing development. GenAI enhances writing in genres such as creative writing (Laksanasut, 2024) and hortatory exposition texts (Amyatun & Kholis, 2023). It has increased learner engagement and motivation (Zakaria & Ningrum, 2023; Song, 2023) while fostering positive perceptions of the writing process (Abdul Rahman et al., 2022). Additionally, it provides adaptive feedback (Ghorbandordinejad & Kenshinbay, 2024), improving learners' writing at the lexical and syntactic levels (Al Mahmud, 2023) and enhancing vocabulary (Austria et al., 2022). GenAI tools also contribute to the coherence and cohesion of texts (Alzahrani & Alotaibi, 2024). By supporting idea generation, improving critical thinking, preventing plagiarism, and raising the quality of academic writing (Bakri et al., 2024), GenAI aids learners in understanding grammatical rules (Han et al.,2021; Wulandari et al., 2024) and encourages self-regulated learning strategies (Wei, 2023). Furthermore, GenAI influences learners' behavior and acceptance of technology (Nazari et al., 2021), improves content organization (Marzuki et al., 2023), and provides scaffolding throughout the writing process (Dhillon et al., 2023). In this research area, scholars have explored diverse foci and methods to address critical questions as to what, why, how, and to what extent GenAI enhances learners' writing competence. With the expanding role of GenAI in writing, a systematic review of related studies is crucial for understanding its nuanced impact on ELT/L.

A systematic review by Chen et al. (2022) has broadly explored GenAI's influence on education, including its trends and challenges. However, they need to specifically focus on ELT/L or writing competence, highlighting a gap in understanding GenAI's impact. This gap emphasizes the need for a focused review of GenAI's role in enhancing writing skills

within ELT/L contexts to guide research and practice better. However, in their recent research article, Crompton et al.(2024) conducted a systematic review of AI in ELT/L, identifying five key areas where AI is utilized: speaking, writing, reading, pedagogy, and self-regulation. While the review highlighted AI's significant impact on vocabulary learning and grammar in writing, the discussion focused on concerns about overreliance on AI tools and a narrow focus on feedback through grammar and spelling checkers.

While previous systematic reviews on the implications of artificial intelligence in ELT/L or education have provided valuable insights, discussions have been mainly limited to aspects such as linguistic accuracy and the ethical implications of AI tools. However, this field still needs to be explored, such as how learners interact with feedback, comments, and suggestions from these tools, particularly across different stages of the writing process. This micro-level view of the relationship between students and AI-DWTs requires an in-depth investigation to capture the nuanced impact of these tools on writing competence. This study thus underscores the need for a comprehensive systematic review to examine research practices in this underexplored area. This review aims to examine previous research to identify various research designs, methods, theories, and perspectives that will facilitate the emergence of best research practices in the field, which other researchers can utilize. In addition, the review aims to address identified methodological or theoretical gaps that future studies can work to fill.

In essence, this review aims to explore the designs, methods, theories, models, and perspectives used in research on AI-DWTs for developing writing competence. This will benefit researchers and scholars who intend to investigate the role of AI-DWTs in developing learners' writing competence. In this context, the following research questions (RQs) arise:

RQ1: What research designs are used in studies on AI-DWTs for teaching and learning writing?
RQ2: What research methodologies, including data collection and mode of analysis, are employed in studies on AI-powered digital writing tools in teaching and learning writing?
RQ3: What theories/models are commonly used in research on AI-DWTs for teaching and learning writing?

2. Methods

This systematic review followed the PRISMA guidelines (Page et al., 2021) and involved two main stages: the search, identification, and selection of relevant studies and the subsequent analysis using mixed methods. Quantitative methods provided aggregate data, while qualitative techniques, including inductive and deductive coding, summarized GenAI's role in ELT/L and generated emerging theories from the collective review (Gough et al., 2019). In parallel, a theoretical review examined the conceptual frameworks and models underpinning current research on AI in ELT/L, assessing their relevance and identifying areas for further

development (Torraco, 2016). Additionally, a methodological review summarized the state-of-the-art research practices in this field, focusing on the design, execution, and synthesis of studies (Chong et al., 2024).

Search strategy

This review analyzed studies from 2022 until the second quarter of 2024 on AI-DWTs in ELT/L (specifically on writing competence), stressing the increased AI use in education following ChatGPT's release in November 2022. Google Scholar was used for its accessibility and valuable features like reference exporting and full-text access (Mikki, 2009). Several Boolean search items were used to explore the literature on AI-DWTs and their impact on writing competence and skills in ELT/L. The first search terms are "AI-powered digital writing tools" OR "AI writing tools") Moreover, "writing competence" AND ("English language teaching" OR ELT). The second search terms are "AI-driven writing tools" OR "digital writing aids") AND "writing skills" AND ("language learning/teaching" OR TMLL). Finally, the third search terms, ("AI-powered writing tools" OR "AI-DWTs") AND ("writing skills" OR "composition skills") AND ("feedback" OR "assessment" OR "evaluation") AND ("English as a Second Language" OR ESL). These search terms collectively facilitate a targeted search for studies on AI-DWTs and writing competence in language education.

The electronic and hand search identified 87 papers for possible inclusion in the systematic review. The 87 were then examined for a match against the inclusion/ exclusion criteria listed in Table 1.

Table 1
Inclusion and Exclusion Criteria

Inclusion	Exclusion
• Journal papers published between 2022 and first quarter of 2024 • Peer-reviewed journal papers	• Editorials • Research that includes students learning about AI, such as a computer science class and not using AI for learning
• Involves writing competence development in English as an ESL and EFL • Use AI-DWTs in developing writing competence • Journal papers written in English	• Teaching and learning English as their L1

The researchers independently reviewed each paper against the criteria, and the interrater reliability was calculated using percentage agreement (Belur et al., 2021). A total of 53 papers were removed after applying the inclusion and exclusion criteria. At the end of the search selection process, 34 papers met the inclusion criteria for this study.

Data analysis

Once the papers were screened and selected, data were extracted to align with the research questions guiding this study. A combination of *priori* and *grounded* coding was used to analyze the data. A priori coding, which involves using predetermined codes before examining the data, was applied to identify common methodological approaches based on the existing literature review. Grounded coding, on the other hand, was applied to identify theoretical approaches inductively from the data to allow new perspectives to emerge, particularly in the new-fangled context of the GenAI era.

3. Results and Discussion

The systematic review presents the synthesis of the identified research designs and methods, including (a) data collection techniques, (b) modes of analysis, (c) theories/models or perspectives used in 34 studies on AI-DWTs for developing writing competence among learners.

3.1 Research designs used in studies on AI-DWTs for teaching and learning writing

The analysis of research designs used in studies on AI-DWTs for teaching and learning writing reveals diverse perspectives as shown in Table 2.

Table 2
Research Designs Used in the Reviewed Studies

Research designs	Count	Percentage (%)
Experimental Design	14	41.18
(Quasi-Experimental)	8	23.53
(True-Experimental)	6	17.65
Survey Design	10	29.41
Systematic Literature Review	5	14.71
Design-Based Research	1	2.94
Action Research (Classroom Action Research)	1	2.94
Case Study	1	2.94
Technology Review	2	5.88
Total	***34***	***100***

The experimental design emerged as the most prevalent in this research area, with 41.18% of the reviewed studies. Within this, quasi-experimental designs with 23.53%, including studies by Laksanasut (2024), Al Mahmud (2023), Tedjo (2022), Liu et al.(2024), Vo and Nguyen (2022), and Mahapatra (2024). On the other hand, true experimental designs with 17.65%, were conducted by Dong (2023), Abdul Rahman et al. 2022), Roisah et al.(2024), Song and Song (2023), Wei (2023), Nazari et al.(2021), lastly Alzahrani and Alotaibi (2024). True experimental and quasi-experimental designs are both valuable for examining cause-and-effect relationships, but they differ in the level of control they offer and the strength of causal conclusions they can support. Hence, selecting the appropriate experimental design is essential, especially in complex research areas like AI-assisted writing development. True experimental design is assigned through randomized methods, which can provide stronger control over variables and participants. This is more suitable for establishing clear causal links. On the other hand, quasi-experimental designs are often more feasible in natural settings where random assignment is impractical.

In essence, these studies that utilized experimental research design were able to measure comparison of the writing skills before and after an intervention, as well as comparisons between traditional teaching methods and those that integrate AI tools like ChatGPT, Grammarly, Wordtune, Quillbot, ProofWriter, and AI-KIKU. It can be noted that comparing traditional teaching methods with approaches that integrate AI tools, such as ChatGPT, Grammarly, Wordtune, Quillbot, ProofWriter, etc., remains common in the reviewed studies. Nevertheless, scholars increasingly question the validity of these comparisons, as they assume that technology in learning environments should align with «traditional» learning conditions. In reality, technology-mediated language learning environments may represent a fundamentally different experience. It offers unique affordances that call for distinct evaluation methods (Madyarov, 2009). This situation presents opportunities for innovative methodologies and perspectives essential to fully explore these differences and assess AI tools on their own merits.

In addition, most of these studies have employed a pretest-posttest control group design, where both the experimental and control groups are pretested before the intervention, followed by a post-test after the intervention. This research design is often preferred over more complex designs, such as the Solomon four-group design and factorial design. The Solomon four-group design and factorial design are advanced experimental designs that can provide valuable insights into investigating AI-DWTs in writing classrooms. The Solomon Four-Group Design controls for potential pretesting effects by including groups that receive pretests and others that do not, allowing researchers to isolate the true impact of AI-DWT interventions. In contrast, the Factorial Design enables the examination of multiple variables, such as tool type, level of scaffolding, learner proficiency, and their interactions. This approach can reveal how different combinations of factors influence writing outcomes, which can provide a nuanced view of how AI-DWTs may best support diverse learners.

The use of survey design was the next most common with 29.41% of the studies, including work by Ananda and Salmiah (2024), Austria et al.(2022), Bakri et al. (2024), Burkhard (2022), Ghafar et al. (2023), Kaharuddin et al. (2024), Le (2023), Ou et al. (2024),

Viktorivna et al. (2022), Zebua and Katemba (2024). Systematic literature reviews were utilized in 14.71% of the studies, such as those by Dergaa et al. (2023), Fazal et al. (2024), Ghorbandordinejad and Kenshinbay(2024), and Zakaria and Ningrum (2023).

Other research designs included design-based research (2.94%) by Tseng and Warschauer (2023), action research (2.94%) by Amyatun and Kholis (2023), case study (2.94%) by Gayed et al.(2022), and technology reviews (5.88%) by Fitria (2023) and Barrot (2023). Although fewer in number than experimental designs, these studies provide valuable qualitative insights into this research area. It can also be noted here that some studies need a clearly defined methods section, which makes it difficult to determine the study's direction and interpret its implications. The absence of a detailed methodology in some study limits the understanding of how these technologies were implemented and evaluated. The limitations of this may need to be clarified in order to understand their effectiveness and application in similar contexts.

Given the complexity surrounding learners' writing competence with the use of AI-DWTs, progress is observed with a growing number of experimental studies in this research area. However, this does not imply that other research designs and approaches, such as survey design, systematic literature reviews, design-based research, action research, case studies, and technology reviews, are less critical. These methodologies can offer valuable insights and contribute a more comprehensive understanding of how AI-DWTs influence writing development.

3.2 Research methods used in studies on AI-DWTs for teaching and learning writing

The review of research methods used in 34 studies on AI-DWTs for teaching and learning writing reveals a diverse methodological landscape as shown in Table 3.

Table 3
Research Methods Used in the Reviewed Studies

Research Methods	Count	Percentage (%)
Quantitative	10	29.41
Qualitative	8	23.53
Mixed-methods	5	14.71
Descriptive Research	*10*	*29.41*
Descriptive qualitative	4	11.76
Descriptive quantitative	6	17.65
N/A	1	2.94
Total	***34***	***100***

The quantitative method is the most prevalent among the reviewed studies, utilized in 10 studies (29.41%). These studies have focused on measurable outcomes, such as improvements in writing skills and engagement with AI-DWTs. Second, qualitative methods were utilized in eight studies (23.53%). These studies have explored learners' personal experiences and perceptions regarding using AI in writing. A more comprehensive method, such as a mixed method, was used in five studies (14.71%). These studies have combined quantitative and qualitative data to understand the impact of AI-DWTs in writing.

Notably, descriptive research was employed in 10 studies (29.41%), with a distribution between descriptive quantitative (17.65%) and descriptive qualitative (11.76%) approaches. Descriptive research is used to summarize trends and patterns without testing specific hypotheses. This offers foundational insights into the role of AI in writing and often serves as the starting point for further research. In essence, descriptive quantitative studies focus on aggregating numerical data to provide a broad overview of trends, while descriptive qualitative studies summarize qualitative data to highlight patterns and themes.

In specific, quantitative studies have been conducted by authors such as Laksanasut (2024), Alzahrani and Alotaibi (2024), and others, while mixed-methods approaches are utilized by researchers like Al Mahmud (2023) and Song and Song (2023). This approach allows for a richer understanding of the data by capturing both numerical trends and contextual insights, which can inform better instructional practices. Interestingly, Liu et al.(2024) focus on qualitative methods to explore digital multimodal composing. The study by Tseng and Warschauer (2023), a design-based study, did not require a specific research method.

Although experimental design is prominent among these studies, most rely solely on quantitative data, particularly test scores, to assess outcomes. Only five studies employed mixed-methods approaches, but these studies lack specification regarding the type of mixed-methods design used—such as convergent, explanatory sequential, or exploratory mixed-methods designs. Clearly specifying the type of mixed-methods approach could provide in-depth insights into the complexities of writing development.

By incorporating qualitative data alongside quantitative metrics, researchers can gain a richer understanding of learners› experiences and behaviors, ultimately enhancing the understanding of how these tools may impact writing competence.

3.2.1 Data collection techniques used in the reviewed studies

The analysis of data collection techniques utilized in 34 studies on AI-driven digital writing tools (AI-DWTs) for teaching and learning writing reveals a rich and varied methodological landscape, as illustrated in Table 4.

Table 4
Data Collection Techniques Used in the Reviewed Studies

Techniques	Count	Percentage (%)
Pre-test and post-test	**12**	35.29
Questionnaire	9	26.47
Data extraction	4	11.76
Semi-structured interview	3	8.82
Focus group discussion	1	2.94
Screen recordings with think-aloud protocol and post project interview	1	2.94
Observation	2	5.88
Examination of library resources	1	2.94
Pre/post-tests with qualitative interviews	1	2.94
Total	***34***	***100***

In the reviewed studies, a variety of data collection techniques were employed, with pre-tests and post-tests being the most prevalent, used in 12 studies (35.29 %). Questionnaires were followed closely and utilized in nine studies (26.47%) of the methods. Data extraction was employed in four studies (11.76%), while semi-structured interviews were used in three studies (8.82%). Other methods included observation (2.94%), focus group discussions (2.94%), and screen recordings with think-aloud protocols (2.94%.). Additionally, pre/post-tests with qualitative interviews appeared in one study (2.94%), as did the examination of library resources. Notably, the study of Liu et al. (2024) utilized screen recordings with think-aloud protocol and post-project interviews as their techniques. The use of data extraction from secondary sources was also accounted for in the review.

Most of these experimental studies anchor their analysis in test scores from the pre-test, post-test, and delayed post-test, whether within or between groups. Nevertheless, investigations into the learning behaviors regarding how students utilize these tools in writing are limited despite the fact that writing involves various recursive processes. Investigating learning behavior, rather than solely relying on test scores, is crucial. While test scores offer a snapshot of performance, they often overlook the complexities of the writing process. Understanding learning behaviors reveals the metacognitive strategies and social processes that underpin language acquisition, such as how students navigate AI-DWTs' resources and functionality and how these factors interact with one another in the stages of writing. This may be fulfilled by capturing screen recordings as learners prompt and respond to feedback strategically. Such data may also be gathered from conversations in the *chat history* of these AI-DWTs.

The review also noted that most studies collected writing outputs after learners used AI-DWTs during individual writing tasks (though most studies did not specify whether tasks

were done individually or in groups). This suggests a limited exploration of AI-DWT usage in collaborative settings. In their study on designing learning in an AI-driven world, Carvalho et al. propose the Activity-Centered Analysis and Design (ACAD) framework (Goodyear & Carvalho, 2014; Goodyear et al., 2021), which provides both theoretical and practical insights for effective learning design. ACAD emphasizes that learning activity is shaped by (a) proposed tasks (epistemic design), (b) available physical and digital tools (set design), and (c) pre-planned social arrangements (social design). The framework centers on the idea that learning activity is emergent and influenced by these designable elements while learners retain the agency to shape their interactions and outcomes.

3.2.2 Mode of analysis used in studies on AI-DWTs for teaching and learning writing

The mode of analysis used in the 34 studies on AI-DWTs for teaching and learning writing reveals a diverse methodological landscape as shown in Table 5.

Table 5
Mode of Analysis Used in the Reviewed Studies

Mode of Analysis	Count	Percentage (%)
Thematic analysis	11	32.35
Frequency distribution	3	8.82
Paired sample t-tests	3	8.82
Independent-sample t-test	4	11.76
Document analysis	2	5.88
Analysis of covariance	4	11.76
Path Analysis	1	2.94
Intraclass Coefficient Correlation	1	2.94
Standard deviation	1	2.94
Percentages, median, mean scores	3	8.82
N/A:	1	2.94
Total	***34***	***100***

The most frequently utilized mode of analysis was thematic analysis, employed by 32.35% of the 34 reviewed studies. Frequency distribution, paired sample t-tests, and percentages, median, and mean scores each accounted for 8.82% of the studies. Additionally, independent-sample t-tests were used with 11.76%, along with an analysis of covariance, which also received the same percentage. Document analysis was less common, utilized by 5.88%, while path analysis, intraclass coefficient correlation, and standard deviation represented 2.94% of the methods used. The study by Tseng and Warschauer (2023) was done without specific data collection or analysis techniques.

Although experimental design and the use of quantitative data are the most prevalent approaches among these studies, the combination of systematic review and survey design of the reviewed studies is actually more common than experimental designs relying solely on quantitative data, which necessitates thematic analysis. Specifically, five of the experimental studies used mixed methods, whereas researchers incorporated pre-test and post-test measures along with questionnaires and interviews, which required thematic analysis. Notably, the study by Liu, et al. (2024), although experimental in nature, necessitated thematic analysis as it drew data from screen recordings, a think-aloud protocol, and post-project interviews.

Thematic analysis was primarily employed in studies using survey designs to gather insights regarding learners' perceptions. However, its application to analyze learners' writing outputs during experimental studies was notably absent in the reviewed studies. Thematizing learners' essays is crucial, as it provides deeper insights into their cognitive processes, understanding of language concepts, and writing development. By identifying recurring themes and patterns in writing, educators can better tailor instructional strategies, enhance feedback mechanisms, and support learners in overcoming specific challenges, ultimately fostering improved writing competence. It can also be noted here that while research on perceptions provides essential insights into user attitudes and potential barriers to adoption, it must be complemented by empirical studies measuring the actual impact of AI-DWTs on writing development for a well-rounded understanding.

3.3 Theories, models, and perspectives in studies on AI-DWTs for developing writing competence

The study identified diverse theories, models, and perspectives in research on AI-DWTs for developing writing competence, as shown in Table 6. This variety underscores the complexity of writing education and the different perspectives researchers use to analyze the implications of AI.

Table 6
Theories, Models, Perspectives Used in the Reviewed Studies

Theory/Approach	Count	Percentage (%)
Technology-Mediated Language Learning (implicitly)	11	32.35
AI in ELT/L	5	14.71
AI Literacy	2	5.88
Social Constructivism	2	5.88
Automated Writing Evaluation (AWE)	2	5.88
Digital Writing in AI-era	2	5.88
Communicative Language Teaching (CLT) approach	1	2.94

Table 6 continued...

Theory/Approach	Count	Percentage (%)
Adaptive Feedback in CALL for Writing Proficiency	1	2.94
Cognitive and Social Constructivist Theories (Krashen, 1982)	1	2.94
Input Hypothesis (Krashen, 1986)	1	2.94
Teaching Writing as Reflective Practice (Hillocks Jr., 1995)	1	2.94
Written Corrective Feedback (WCF)	1	2.94
Construct of One's Identity (Ivanic, 1998)	1	2.94
Communicative Competence (CC)	1	2.94
Digital Multimodal Composing	1	2.94
Fowler (2007) Writer's Perception and Exploration	1	2.94
Technology Affordances	1	2.94
Scaffolding in Writing	1	2.94
Total	*34*	*100*

The reviewed studies utilize various perspectives to explore the use of AI-DWTs in writing development. Laksanasut (2024) applies the communicative teaching approach, which focuses on improving the language skills of the students to communicate effectively and meaningfully in real-world situations. Bakri et al. (2024) draw on Stephen Krashen›s Input Hypothesis (1986) and George Hillocks Jr.›s «Teaching Writing as Reflective Practice» (1995), which focuses on language acquisition through comprehensible input and reflective practice for writing improvement. Song and Song (2023) use Vygotsky's Social Constructivism, which emphasizes social interaction and the Zone of Proximal Development (ZPD) in learning. Wei (2023) also employs Vygotsky's Social Constructivist Theory, which highlights the role of social interaction and scaffolding in writing development through AI. Fazal et al. (2024) use the Technology Acceptance Model (TAM) to examine factors influencing the adoption of AI tools in writing. Lastly, Dhillon et al. (2023) bank on scaffolding. They found that different levels of scaffolding from large language models (LLMs) influenced the co-writing process. In their study, participants responded to argumentative writing prompts under three conditions: (a) no AI assistance, (b) next-sentence suggestions (low scaffolding), and (c) next-paragraph suggestions (high scaffolding). Other perspectives highlight the growing focus on AI's relevance to English Language Teaching and Learning (ELT/L) (14.71%) and AI literacy (5.88%). Despite the various perspectives that these studies were anchored, a significant proportion of the reviewed studies lack a solid theoretical foundation (32.35%) but seem to lean toward perspectives aligned with TMLL.

Identifying the theoretical background requires careful consideration since many studies lack an explicit theoretical framework. The researchers have to check the background of the study to gain insights into their theoretical framework. With this dilemma, lacking a solid theoretical foundation, it is safe to say that generalizing a clear theoretical trend in this research area may be challenging.

Nevertheless, the analysis categorized the studies using wider theoretical umbrellas, such as the cognitive and social orientations. This was done by looking at the approaches taken, such as how learners used the AI-DWTs in their writing tasks (individually, in pairs, or groups). Table 7 summarizes the theoretical anchors in the 34 studies from 2022 to 2024.

Table 7
Distribution of Theoretical Approaches Used in the Reviewed Studies

Year	Cognitive	Socially-oriented	No clear/relevant background	Total
2022	1	0	4	**5**
2023	5	3	4	**12**
2024	8	2	3	**13**
Total	14	5	11	**34**

Cognitive-informed frameworks dominate the literature, with 14 studies (41%) emphasizing individual learning processes. In contrast, only 5 studies (15%) adopt socially-oriented perspectives, suggesting that social dynamics in learning remain underexplored. Additionally, 11 studies (32%) fall into the "unclear/relevant background" category, highlighting the need for clearer theoretical frameworks. Cerezo (2015) argues that effectively investigating how technology facilitates learning requires a theoretical paradigm. Therefore, to better understand the relationship between learners and AI-DWTs, research must integrate and interpret TMLL practices within a robust theoretical framework. This approach will offer a more comprehensive understanding of the role of AI-DWTs in writing development. In Cerezo's (2015) systematic review of theoretical approaches in CALL research, he found that 67 out of 115 empirical studies (58.3%) employed either a cognitive or socially-oriented theoretical framework, while 48 studies (41.7%) lacked a specific theoretical foundation. This underscores the persistent challenge of insufficient theoretical grounding in research on TMLL, even in the GenAI era.

4. Conclusions and Recommendations

This review sheds light on the various research designs and methods, including data collection techniques and modes of analysis, which researchers may use in investigating the role of AI-DWTs in developing writing competence. While this study reveals methodological

and theoretical opportunities in investigating this research area, challenges also emerge. Experimental designs have yielded valuable insights, but an overreliance on quantitative methods, primarily on gathering and analyzing test scores, limits understanding the nuanced processes underlying learners' engagement and writing development. A lack of direct observation of learners' behaviors and experiences during writing tasks may be bridged by future studies by exploring other types of experimental designs, such as the Solomon four-group design and factorial design, that can provide multiple avenues in investigating the impact of AI-DWTs in writing.

Given the prevalence of studies relying on quantitative data, future research could integrate qualitative data through mixed-methods approaches to achieve a more comprehensive understanding of learners' experiences and behaviors. Researchers may consider several mixed-methods designs, such as the *convergent parallel design*, which simultaneously collects and analyzes both data types to compare and integrate results; the *explanatory sequential design*, which begins with quantitative data collection and is followed by qualitative data to further explain the findings; and the *exploratory sequential design*, which starts with qualitative insights that are subsequently tested with quantitative methods. Other designs include the *embedded design*, where one data type serves as primary and the other as supplementary for additional context, and the *multiphase design*, which integrates multiple mixed-methods studies over time to address broader research objectives.

More innovative data collection methods are needed to gain a fuller understanding of learners' interactions with AI-DWTs, moving beyond the limitations of questionnaire-based studies. Techniques like screen recordings, think-aloud protocols, and learner reflections can provide real-time insights into the nuanced and complex learning behaviors that AI tools influence, going beyond accuracy-focused metrics. Content analysis of student outputs produced with AI-DWTs can reveal how AI-driven syntax and lexis choices impact language use, while thematic analysis of learner-AI interactions can uncover specific writing strategies, such as prompting, accepting, and responding to AI suggestions. This approach enables researchers to assess not just task outcomes but also the process and quality of interactions, thereby highlighting how AI affordances shape writing strategies in ways that traditional metrics may overlook.

The prevalence of cognitive frameworks over socially oriented ones highlights the need to explore collaborative dynamics in AI-enhanced learning environments further. The lack of clear theoretical foundations in many studies remains a persistent issue that hinders the consistency and replicability of insights. To better understand the collaborative potential of AI-DWTs, future research may incorporate socially oriented theoretical frameworks, shedding light on group dynamics, peer learning, and how AI tools influence collaboration, feedback exchange, and mutual support among learners. Additionally, more defined theoretical frameworks are essential for ensuring consistency and depth in this research field. Researchers should aim to integrate cognitive, social, and interactional frameworks, such as Activity-Centered Analysis and Design (ACAD).

With the emergence of AI-driven learning contexts, there is a growing need for frameworks that can thoroughly capture the dynamics of learner-AI interactions, making

frameworks like the Learner-Computer Interaction (LCI) by Caws and Hamel (2016), which focuses on understanding how learners engage with and interact with computer-based systems. In the GenAI (Generative AI) era, incorporating Learner-Computer Interaction (LCI) becomes even more relevant as AI tools like ChatGPT, Grammarly, and other writing assistants become integral to learning environments. The LCI framework can be applied in this field by examining how learners interact with these AI tools and how their behaviors, cognitive processes, and learning outcomes are influenced by AI-generated content and feedback. This framework focuses on three key processes: task identification, strategy analysis, and usability testing. Task identification pinpoints specific language tasks learners must perform and the patterns involved. Strategy analysis examines the behavioral and metacognitive strategies that learners use to accomplish these tasks effectively, including how learners prompt and process information. Lastly, usability testing assesses the efficiency and user satisfaction of language learning systems to ensure they meet learners' needs.

The ongoing investigation of AI-DWTs in writing development presents considerable potential for advancing educational landscapes. Exploring the range of research methodologies, incorporating creative data collection techniques, and integrating more robust theoretical frameworks will enable future studies to offer more profound, more nuanced insights into the intricate dynamics of learner interactions with AI-DWTs. As the field progresses, it is essential to remain attuned to the evolving landscape of AI in writing instructions, that is, to ensure that these technologies contribute meaningfully to developing writing competence among students.

References

Abdul Rahman, N. A., Zulkornain, L. H., & Hamzah, N. H. (2022). Exploring artificial intelligence using automated writing evaluation for writing skills. *Environment Behaviour Proceedings Journal, 7*(SI9), 547–553. https://doi.org/10.21834/ebpj.v7iSI9.4304

Al Mahmud, F. (2023). Investigating EFL students' writing skills through artificial intelligence: Wordtune application as a tool. *Journal of Language Teaching and Research, 14*(5), 1395–1404. https://doi.org/10.17507/jltr.1405.28

Alzahrani, F. K., & Alotaibi, H. (2024). The impact of artificial intelligence on enhancing EFL writing skills among high school students. *Journal of Educational and Human Sciences, 34*, 226–240. https://doi.org/10.33193/JEAHS.34.2024.482

Amyatun, R. L., & Kholis, A. (2023). Can artificial intelligence (AI) like QuillBot AI assist students' writing skills? Assisting learning to write texts using AI. *ELE Reviews: English Language Education Reviews, 3*(2), 135–154.

Austria, F. Z., Camagay, E. B. A., Domasing, M., Gelido, A., Libatique, C. V., & Cabutotan, V. M. (2022). Extent of Influence of Using AI Chatbots in the Writing Skills of BSE Students of Pangasinan State University – Lingayen Campus. *Southeast Asian Journal of Science and Technology*, *7*(1), 39-47. https://www.sajst.org/online/index.php/sajst/article/view/273

Bakri, I., Wulandari, M. F., Amalia, S. R., & Rut, W. M. (2024). QuillBot integration in the learning of English writing skills (Perception of business management students). *International Journal of Research on English Teaching and Applied Linguistics*, 5(1). https://doi.org/10.30863/ijretal.v5i1.6412

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing, 57*, Article 100745. https://doi.org/10.1016/j.asw.2023.100745

Belur, J., Tompson, L., Thornton, A., & Simon, M. (2021). Interrater reliability in systematic review methodology: Exploring variation in coder decision-making. *Sociological Methods & Research, 50*(2), 837–865.

Burkhard, M. (2022). Student perceptions of AI-powered writing tools: Towards individualized teaching strategies. In *Proceedings of the 19th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2022)*. Lisbon, Portugal.

Caws, C., & Hamel, M.-J. (Eds.). (2016). *Language-Learner Computer Interactions: Theory, methodology and CALL applications* (Vol. 2). Language Studies, Science and Engineering.

Chong, S. W., Bond, M., & Chalmers, H. (2024). Opening the methodological black box of research synthesis in language education: Where are we now and where are we heading? *Applied Linguistics Review, 15*(4), 1557–1568.

Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13460

Crompton, H., Jones, M. V., & Burke, D. (2022). Affordances and challenges of artificial intelligence in K-12 education: A systematic review. *Journal of Research on Technology in Education*, 1–21. https://doi.org/10.1080/15391523.2022.2121344

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

Dhillon, P. S., Molaei, S., Li, J., Golub, M., Zheng, S., & Robert, L. P. (2024). Shaping human-AI collaboration: Varied scaffolding levels in co-writing with language models. *arXiv*. https://arxiv.org/abs/2402.12345

Dong, Y. (2023). Revolutionizing academic English writing through AI-powered pedagogy: Practical exploration of teaching process and assessment. *Journal of Higher Education Research, 4*(2), 52. https://doi.org/10.32629/jher.v4i2.1188

Eslit, E. (2023, April 5). Concept paper on technology integration and English language teaching in the Philippines context. *ResearchGate*. https://www.researchgate.net/publication/369794062_Concept_paper_on_Technology_Integration_and_English_Language_Teaching_in_the_Philippines_context

Ghorbandordinejad, F., & Kenshinbay, T. (2024). Exploring AI-driven adaptive feedback in the second language writing skills prompt. *Journal of Effective Teaching Methods, 2*(3). https://doi.org/10.59652/jetm.v2i3.264

Fazal, N., Tahir, M. S., Chaudhary, M., & Abbasi, M. (2024). Effectiveness of AI integration into computer-assisted language learning (CALL) on student writing skills based on gender. *Pakistan Journal of Humanities and Social Sciences, 12*(1), 224–230. https://doi.org/10.52131/pjhss.2024.v12i1.1974

Fitria, T. N. (2023). Artificial intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essay. *ELT Forum: Journal of English Language Teaching, 12*(1), 44–58. https://doi.org/10.15294/elt.v12i1.64069

Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing assistant's impact on English language learners. *Computers and Education: Artificial Intelligence, 3*, Article 100055. https://doi.org/10.1016/j.caeai.2022.100055

Ghafar, Z. N., Salh, H. F., Abdulrahim, M. A., Farxha, S. S., Arf, S. F., & Rahim, R. I. (2023). The role of artificial intelligence technology on English language learning: A literature review. *Canadian Journal of Language and Literature Studies, 3*(2), 17–31. https://doi.org/10.53103/cjlls.v3i2.87

Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology, 26*(2), 5–24. http://doi.org/10125/73474

Goodyear, P., Carvalho, L., & Yeoman, P. (2021). Activity-centred analysis and design (ACAD): Core purposes, distinctive qualities and current developments. *Educational Technology Research & Development, 69*, 445–464. https://doi.org/10.1007/s11423-020-09926-7

Goodyear, P., & Carvalho, L. (2014). Framing the analysis of learning network architectures. In L. Carvalho & P. Goodyear (Eds.), *The architecture of productive learning networks* (pp. 48–70). Routledge.

Gough, D., Thomas, J., & Oliver, S. (2019). Clarifying differences between reviews within evidence ecosystems. *Systematic Reviews, 8*, Article 1.

Han, Y., Zhao, S., & Ng, L. L. (2021). How technology tools impact writing performance, lexical complexity, and perceived self-regulated learning strategies in EFL academic writing: A comparative study. *Frontiers in Psychology, 12*, Article 752793. https://doi.org/10.3389/fpsyg.2021.752793

Kaharuddin, A. D., Mardiana, L. I., Arafah, B., & Suryadi, R. (2024). Defining the role of artificial intelligence in improving English writing skills among Indonesian students. *Journal of Language Teaching and Research*, *15* (2), 568-578. https://doi.org/10.17507/jltr.1502.25

Laksanasut, S. (2024). The pedagogical impact of ChatGPT on high-school students' creative writing skills: An exploration of generative AI-assisted writing tools. *Journal of National Educational Testing and Evaluation*, *2567*, 17-34.

Le, P. T. N. (2023). Students' perceptions of the AI technology application in English writing classes. *Proceedings of the 20th AsiaCALL International Conference (AsiaCALL2023), 4*. https://doi.org/10.54855/paic.2344Liu

Liu, C., Zhang, Y., & Biebricher, C. (2024). Investigating students' cognitive processes in generative AI-assisted digital multimodal composing and traditional writing. *Computers & Education, 211*, 104977. https://doi.org/10.1016/j.compedu.2023.104977

Madyarov, I. (2009). Designing a workable framework for evaluating distance language instruction. *CALICO Journal, 26*(2), 209–308.

Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education, 10*(2), Article 2236469. https://doi.org/10.1080/2331186X.2023.2236469

Mikki, S. (2009). Google Scholar compared to Web of Science: A literature review. Nordic *Journal of Information Literacy in Higher Education, 1*(1), 41–51. https://doi.org/10.15845/noril.v1i1.10

Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of artificial intelligence-powered digital writing assistant in higher education: Randomized controlled trial. *Heliyon, 7*(7), e07014. https://doi.org/10.1016/j.heliyon.2021.e07014

Ou, A. W., Stöhr, C., & Malmström, H. (2024). Academic communication with AI-powered language tools in higher education: From a post-humanist perspective. *System, 121*, 103225. https://doi.org/10.1016/j.system.2024.103225

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*.

Roisah, S., Widyaningsih, T. L., & Purwananti, Y. S. (2024). The impact of ChatGPT use on EFL students' writing ability. *Concept: Journal of Social Humanities and Education, 3*(3), 191–199. https://doi.org/10.55606/concept.v3i2.1403

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1260843

Stapleton, P., & Radia, P. (2010). Tech-era L2. *Journal of Language Learning and Technology, 14*(3), 55–68.

Tedjo, E. W. (2022). Providing online feedback using an AI proof-reading tool to enhance student surface-level writing skills. *EBONY Journal of English Language Teaching Linguistics and Literature, 2*(1), 37-45. https://doi.org/10.37304/ebony.v2i1.4040

Tseng, W. & Warschauer, M. (2023). AI-writing tools in education: if you can't beat them, join them. *Journal of China Computer-Assisted Language Learning*, *3*(2), 258-262. https://doi.org/10.1515/jccall-2023-0008

Torraco, R. J. (2016). Writing integrative literature reviews: Using the past and present to explore the future. *Human Resource Development Review, 15*(4), 404–428.

Viktorivna, K. L., Oleksandrovych, V. A., Oleksandrivna, K. I., & Oleksandrivna, K. N. (2022). Artificial intelligence in language learning: What are we afraid of. *Arab World English Journal (AWEJ) Special Issue on CALL, 8*, 262–273. https://dx.doi.org/10.24093/awej/call8.18

Wei, L. (2023). Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology, 14*, 1261955. http://doi.org/10.3389/fpsyg.2023.1261955

Wulandari, F., Astuti, M.T., & Marhamah, M. (2024). Enhancing Writing Literacy Teachers' through AI Development. *Jurnal Onoma: Pendidikan, Bahasa, dan Sastra*.

Zakaria, & Ningrum, S. (2023). ChatGPT's impact: The AI revolution in EFL writing. *Borneo Engineering & Advanced Multidisciplinary International Journal (BEAM), 2* (Special Issue TECHON 2023), 32-37.

Zebua, J. A., & Katemba, C. V. (2024). Students' perceptions of using the OpenAI ChatGPT application in improving writing skills. *Journal of Language and Education, Computer Science*, 4(1), 110-123. https://doi.org/10.36312/jolls.v4i1.1805

Declaration of Use of Generative AI

The authors declare that they used ChatGPT and Grammarly only to identify and rectify spelling and typographical errors.

Nikko J. Domingo is an Assistant Professor at the Nueva Ecija University of Science and Technology, where he teaches Purposive Communication and English for specific professional fields. He is currently writing his dissertation, titled *"Learner-Artificial Intelligence Interactions for Writing Competence Development in AI-Enabled Writing Classrooms: A Linguistic Ergonomics Investigation,"* for the degree Ph.D. in English Language Studies at the University of Santo Tomas in Manila, the Philippines. His research interests include language teaching and learning, language and technology, and forensic linguistics.

Rachelle Ballesteros-Lintao, Ph.D. a full professor at the University of Santo Tomas in Manila, the Philippines, is the incumbent chair of the Department of English, Faculty of Arts and Letters and the program lead of the English Language Studies Cluster in the Graduate School. She is the president of the Linguistic Society of the Philippines and the country representative of Clarity (an international plain language organization). Her research interests include forensic linguistics, language teaching and learning, online discourse, and language and technology. A member of the ImpleMendez, she is involved in a European project on Establishing Networks to Implement the Principles on Effective Interviewing for Investigations.