



Digital Storytelling as a Scaffold for Academic Writing: A Sequential Explanatory Mixed Methods Study in Tertiary EFL Contexts

Abebaw Andargie Melaku *

Department of English Language and Literature, Debre Markos University, Debre Markos, Ethiopia

Received: 03 July 2025

Accepted: 10 September 2025

Published: 08 October 2025

ARTICLE INFO.

Key words/phrases:

Academic writing, Digital storytelling, EFL, Scaffolding, Sequential explanatory mixed methods, Tertiary education

Abstract

Although digital storytelling has attracted considerable scholarly interest as a pedagogical tool in language education, relatively little empirical research has examined its specific role as a scaffold for academic writing in English as a Foreign Language (EFL) higher education institutions. The present study used a sequential explanatory mixed-methods design to examine the effects of a structured digital storytelling intervention on EFL undergraduate academic writing competence and to explore students' perceptions of the mechanisms underlying the observed improvements. Sixty-eight first-year students ($n = 34$ per group) from a public university in the Amhara Regional State of Ethiopia were assigned, through cluster randomization of two intact classes, to a digital storytelling condition or a conventional process-writing condition. Pre- and post-test writing scores were analyzed using Analysis of Covariance (ANCOVA), and two focus-group discussions with purposively selected participants, together with weekly reflective journals, were used to explain the quantitative findings. Quantitative results indicated that the experimental group significantly outperformed the control group on overall writing quality ($F(1, 65) = 18.43, p < .001, \eta^2 = .22$), with the strongest gains in coherence and cohesion, followed by lexical resource and task achievement; the gain in grammatical range and accuracy was not statistically significant. Qualitative findings explained these gains, showing that digital storytelling functioned as a cognitive, multimodal, affective, and social scaffold that helped students transition from personal narrative competence to formal academic discourse. The study demonstrates that digital storytelling is a viable and effective transitional scaffold that connects personal narrative competence with formal academic discourse when it is implemented systematically as a component of an academic writing curriculum. When used in a structured way as an augmentative supplement to the standard writing curriculum, DST can be most effective as a way to engage the affective, cognitive, and social components of writing that might be overlooked in traditional methods, developmentally moving from personal narration to disciplinary argumentation.

1 Introduction

With the globalization of higher education, there has been increasing pressure on undergraduates to write competently in English, especially in situations where English is not a native language. These difficulties are exacerbated in sub-Saharan African universities, where students' native languages are

often diverse, exposure to native texts is minimal, and teaching methods are often more concerned with grammatical correctness than with proficiency at the text level of instruction (Bitchener & Ferris, 2012; Hyland, 2019). Although there has been increasing literature on technology-enhanced language learning, the systematic incorporation of

digital tools as scaffolding to support academic writing development in tertiary EFL contexts is still an under-explored field.

Digital storytelling (DST), which is a pedagogically rich medium in which language skills are developed by making short, personal films that combine voice, image, and text (Robin, 2016), has been proposed as a medium for developing language skills. Those who advocate for DST point to its multimodal, narrative, and audience-oriented qualities as stimulating cognitive and affective aspects of composing that are largely unengaged in traditional writing instruction (Lambert, 2018; Ohler, 2013). However, research that directly links DST to formal academic writing products is scarce, and research beyond mere description of students' experiences to account for this connection is even more limited.

In many public universities in Ethiopia, first-year students enter degree programs without a good understanding of the conventions of academic writing in English (Fantaye, 2020; Mulugeta & Getnet, 2022). Product- and process-oriented approaches remain the dominant instructional models in most existing writing courses in this context; however, despite their pedagogical value, they offer few opportunities for students to develop the rhetorical awareness and authorial voice that disciplinary participation requires (Lea & Street, 2006; Lillis & Curry, 2010). These universities are slowly becoming digital, but digital technologies are rarely used purposefully as a pedagogical scaffold. Existing Ethiopian studies have documented that these writing difficulties persist despite curricular reform (Fantaye, 2020; Yimam, 2018; Mulugeta & Getnet, 2022); yet almost none of this literature has examined whether a structured, multimodal intervention such as DST can function as a bridge between students' existing narrative competence and the unfamiliar demands of the academic register. It is this practical gap—the absence of an empirically validated, low-cost scaffold suited to large, under-resourced EFL writing classes—together with the theoretical gap identified above, that motivates the present inquiry.

Given the practical and theoretical significance of this question, the present study used a sequential explanatory mixed-methods design (Creswell &

Plano Clark, 2018) to pursue the following three objectives.

The quantitative strand investigated the impact of a structured DST intervention on the academic writing skills of first-year EFL undergraduates, whereas the qualitative strand involved student reflective journals and focus-group discussions to explain the patterns observed in students' performance. This integration of measurement and meaning-making positions the study to offer evidence both for the effect of DST and for a theory of how DST works as a scaffold.

Three research questions guided the inquiry:

- Is there a statistically significant difference in academic writing performance between EFL undergraduates who participate in a structured digital storytelling intervention and those who receive conventional writing instruction?
- Which sub-constructs of academic writing are most significantly affected by the digital storytelling intervention?
- How do students perceive digital storytelling as influencing their academic writing processes and outcomes?

2 Review of Related Literature

2.1 Scaffolding Theory and Academic Writing

The “zone of proximal development” (ZPD) is a term used in education to refer to the scaffolding needed to help learners achieve something they cannot do on their own, as developed by Vygotsky (1978). Within the field of writing pedagogy, scaffolding is implemented in various ways, such as through the use of peer feedback, genre modelling, process writing conferences, and, more recently, digital tools (Gibbons, 2015; Hammond, 2020). Following a distinction made by Hammond and Gibbons (2005) between macro-scaffolding (deliberate design of curriculum) and micro-scaffolding (moment-to-moment instructional support), it is important to consider how DST works pedagogically. The scaffolding metaphor suggests a temporary state: good scaffolds gradually fade as learners'

skills increase. This is the theoretical inquiry that drives the present study: Can DST play this role in helping students prepare for academic discourse?

In the present study, this distinction is operationalized directly: the four sequenced DST cycles described in Section 3.3 constitute macro-scaffolding, in that the curriculum was deliberately pre-planned to move learners from low-stakes personal narration toward increasingly disciplinary, argumentative, and finally formal academic registers, with support incrementally withdrawn as each cycle's cognitive and rhetorical demands increased. This macro-level design was complemented by micro-scaffolding in the form of instructor modelling and peer feedback offered within each cycle. Framing the intervention in these terms clarifies that it is the structured progression of cycles—not digital storytelling as a generic activity—that constitutes the scaffold examined in this study.

2.2 Academic Writing in EFL Tertiary Contexts

Academic writing can be defined as a register-specific, socially situated form of literacy characterized by explicit argumentation, hedged claims, disciplinary citation conventions, and a structured, reader-oriented organization of ideas (Hyland, 2019). Studies in the context of Ethiopian higher education reveal continued challenges with discourse organization, hedging, citation conventions, and personal voice negotiation in the academic register (Fantaye, 2020; Yimam, 2018). These are not only linguistic but also epistemological challenges, because learners need to know how to become knowledge constructors in disciplinary communities where the rules of engagement are often hidden and presented as other cultures' norms rather than made explicit (Lillis & Curry, 2010).

Thus, the theoretical framing of Academic Literacies (Lea & Street, 1998; 2006) is relevant. Academic Literacies emphasizes the identity-related and context-specific nature of writing development, rather than emphasizing the transferability of academic skills. This framing offers an opportunity to reflect on the potential of DST as a genre that emphasises personal voice and authentic communication as a means of supporting identity construction in the

process of positioning students as valid academic writers.

Taken together, this literature indicates that academic writing development in EFL tertiary contexts is not solely a matter of grammatical accuracy but also one of identity and rhetorical socialization—a gap that a narrative-based, identity-affirming intervention such as DST is theoretically well positioned to address.

2.3 Digital Storytelling in Language Education

Digital storytelling made its significant presence in language education literature in the early 2000s, when the Centre for Digital Storytelling (currently Story Center) was founded by Joe Lambert and his colleagues (Lambert, 2018). Initial applications were in the area of L1 literacy and affective aspects of learning, after which the focus shifted to L2 and foreign language contexts of DST (Sylvester & Greenidge, 2009; Vinogradova *et al.*, 2011). Although there have been positive findings regarding general language proficiency, oral fluency, and vocabulary acquisition, writing-specific outcomes, especially formal academic writing, have been less systematically studied, as indicated by the results of meta-analytic reviews conducted by Nguyen (2021).

Concerning the present study, the work of Abdel-Hack and Helwa (2014) on an Egyptian university student population that showed significant improvements in their English writing proficiency after exposure to DST is relevant, as is that of Castañeda (2013), who used DST to enhance thesis-driven writing among case study participants in a community college in the United States. However, most of these studies happen to be in Western or first-language contexts, and the processes by which DST facilitates the development of writing were rarely systematically investigated, examined, and questioned.

The benefits that have been reported as a result of DST in language education are as follows: it improves learners' motivation and affective investment, as it gives them the chance to develop their own voice and voice control through authentic authorship (Lambert, 2018); it offers opportunities

for iterative revision across various modes, which leads to metalinguistic awareness (Vinogradova *et al.*, 2011); and it provides a low-stakes rehearsal space for learners to practice voice and structure before transferring them to the higher-stakes setting of academic tasks (Sylvester & Greenidge, 2009). Regarding writing quality specifically, the recent empirical study by Castillo-Cuesta *et al.* (2021) with 101 EFL pre-service teachers confirmed these results, with a significant improvement in their writing performance after the DST intervention using Storybird. Similarly, Ajabshir (2024), who adopted a quasi-experimental design with complexity, accuracy, lexical diversity, and fluency (CALF) measures, revealed that DST significantly increased EFL learners' syntactic complexity, accuracy, and lexical diversity compared to a traditional storytelling comparison group, but did not significantly influence their fluency, highlighting that DST has specific value for the more formally oriented aspects of writing that correspond to the academic quality of writing. These benefits provide the theoretical basis for expecting that a cycle-based DST intervention could function as the kind of macro-scaffold described in Section 2.1.

2.4 The Narrative-Academic Writing Interface

There are a number of theorists who have suggested that narrative competence itself is a developmental resource for academic writing (Bruner, 1990; Zinsser, 2013). Narrative thinking involves temporal sequencing, character motivation, and causal reasoning, all of which are common to argumentative and analytical writing. At the tertiary level, Dyson and Genishi (1994) showed how children's oral language supported their learning of literacy, and similar mechanisms may be at work here when students rehearse the structure of academic arguments through a narrative format. The narrative-analytical interface may be activated in ways different from traditional instruction, as students are asked to create a visual and verbal narrative before they translate their understandings into written academic prose in DST.

Together, Sections 2.1–2.4 establish the theoretical rationale for the present study: scaffolding theory explains how a structured DST sequence could support writing development; the Academic

Literacies and narrative-writing literature explain why personal narrative competence should transfer to the formal academic register; and the empirical DST literature identifies a clear gap—the absence of research that connects DST directly and systematically to formal academic writing outcomes in an under-resourced EFL setting. The present study addresses this gap.

3 Methods

3.1 Research Design

This study used a sequential explanatory mixed-methods design, as described by Creswell and Plano Clark (2018), which involved collecting and analyzing quantitative data first and then collecting qualitative data as a deliberate explanation for the quantitative data. This design is especially appropriate for educational research in which effect size data are not sufficient to make instructional decisions; it is important to know “how” an intervention works as well as whether it is effective.

3.2 Participants and Setting

This study was carried out at Debre Markos University, a public university in the Amhara Regional State of Ethiopia, in the 2023/2024 academic year. The university was selected through convenience sampling on the basis of the researcher's institutional access and the availability of an English for Academic Purposes (EAP) course organized into two parallel intact sections, which made a cluster-based group comparison feasible. The sample consisted of 68 first-year undergraduate students from two intact classes in a mandatory English for Academic Purposes (EAP) course. Because intact class sections rather than individual students were the unit of assignment, one of the two sections ($n = 34$) was randomly assigned to the digital storytelling condition and the other ($n = 34$) to the conventional process-writing condition; this is therefore a cluster-randomized, quasi-experimental design rather than individual-level random assignment, a distinction discussed further as a limitation in Section 5.2. Participants ranged in age from 18 to 24 years ($M = 19.7$, $SD = 1.3$). There was no prior formal exposure to digital storytelling as a pedagogical tool reported by any of the partic-

ipants, and all had finished secondary education with English as the language of instruction.

An analysis for a between-groups ANCOVA with one covariate, assuming a medium effect size ($f = 0.25$), $\alpha = .05$, and power = .80, indicated a required minimum sample size. The achieved sample of 68 was smaller than this benchmark; the study is therefore adequately powered only to detect medium-to-large effects, and this constraint is acknowledged as a limitation in Section 5.2.

In Phase II, 12 participants were purposively selected from the experimental group in accordance with maximum variation sampling (Patton, 2015) to ensure that they represented a variety of performance trajectories, from high- to mid- to low-progress writers (6 women and 6 men).

3.3 Intervention

The DST intervention lasted eight weeks (16 contact hours) and was integrated into the EAP course curriculum. The experimental group went through four digital story cycles, each ending with a 3–5-minute digital story created with free, browser-based software (Canva, Adobe Express). Each cycle was designed around a pedagogical sequence developed along the Seven Elements framework of Lambert (2018) and macro-scaffolding principles described by Gibbons (2015):

- Week 1–2 (Cycle 1): Biographical narratives—students told a story of a major experience in their learning.
- Week 3–4 (Cycle 2): Explanatory narratives—students narrated an explanation of a concept from their academic discipline.
- Weeks 5–6 (Cycle 3): Argumentative narratives—students developed a position-based narrative about a socially relevant issue.
- Week 7–8 (Cycle 4): Academic abstracts—students converted the “plot” of Cycle 3 into a formal academic abstract and introduction.

The last cycle provided a link between personal narrative and formal academic writing. The control

group followed the same EAP curriculum objectives with eight weeks of conventional process-writing instruction, including pre-writing, drafting, peer review, and revision.

Throughout the intervention, the first author—who was also the course instructor—designed and delivered the DST cycles, facilitated in-class story-planning and peer-feedback sessions, and conducted brief formative check-ins after each cycle to confirm task completion and provide individual guidance; these formative check-ins were not scored and did not contribute to the summative pre-/post-test writing measure reported in Section 4.

3.4 Instruments

3.4.1 Academic Writing Test

Academic writing skills were evaluated by a parallel-form extended writing task (500 words) at the pre-test (Week 0) and post-test (Week 9). In both, students were asked to write an argumentative academic essay on one of two counterbalanced, curriculum-relevant prompts, with prompts matched for topic familiarity and difficulty and counterbalanced across groups and testing occasions to control for topic-effect confounds. Essays were marked by two trained readers following an adapted version of the IELTS Academic Writing Task 2 scoring rubric (British Council, 2022), scored in four categories: Task Achievement (TA), Coherence and Cohesion (CC), Lexical Resource (LR), and Grammatical Range and Accuracy (GRA), with each criterion marked on a 0–9 scale. Inter-rater reliability was established before scoring (Cohen's $\kappa = .81$), and differences in band scores of more than 1 point were adjudicated by a third rater.

3.4.2 Reflective Journals

Participants in Phase II kept weekly reflective journals throughout the 8-week intervention. Students were prompted to reflect on how they perceived the relationship between their storytelling work and their academic writing processes. A total of 96 journal entries were collected (12 participants $\times$ 8 weeks).

Journal prompts were reviewed by two EAP specialists for content relevance and clarity before use. Trustworthiness of the journal data was further supported through triangulation with the focus-group data and through member checking, as described in Section 3.5.2.

3.4.3 Focus Group Discussion

In Week 9, after completion of the post-test, two focus-group discussions (FGDs) were conducted with six participants in each group. Conversations were conducted in English, with Amharic used when participants needed linguistic precision, and were audio-recorded and translated. The average duration of the FGDs was approximately 75 minutes. Questions were designed to probe and elaborate on the quantitative findings; for example, they targeted reasons why coherence gains were especially strong in the experimental group.

Whereas the reflective journals captured individual, in-the-moment reflections distributed across the intervention, the FGDs were designed to generate group-level, retrospective elaboration after the intervention had concluded, allowing participants to compare experiences and collectively interpret the quantitative patterns already observed. The two instruments were therefore complementary rather than redundant.

3.5 Data Analysis

3.5.1 Quantitative Analysis

A one-way Analysis of Covariance (ANCOVA) was used to control for pre-test differences in writing ability and to determine post-test differences between groups, with pre-test writing scores serving as a covariate. Four separate ANCOVAs were performed for each of the four sub-constructs and the composite score. Partial eta-squared (η^2) is reported as an effect size. Data were analyzed using SPSS Statistics 27 software.

Before conducting the ANCOVAs, the underlying assumptions were examined. Normality of residuals was assessed using the Shapiro–Wilk test, homogeneity of variance using Levene’s test, and homogeneity of regression slopes by testing the group $\times$ covariate interaction term. There were no significant Shapiro–Wilk test results for the four

sub-constructs in any group (all $p > .05$), suggesting no meaningful deviation from normality. Levene’s test indicated equal variances across groups for Task Achievement, $F(1, 66) = 0.02$, $p = .899$; Coherence and Cohesion, $F(1, 66) = 0.00$, $p = .986$; Lexical Resource, $F(1, 66) = 0.00$, $p = .974$; and Grammatical Range and Accuracy, $F(1, 66) = 0.23$, $p = .635$. The group $\times$ pre-test interaction term was not significant for any of the four sub-construct models (Task Achievement: $F(1, 64) = 1.97$, $p = .165$; Coherence and Cohesion: $F(1, 64) = 0.03$, $p = .869$; Lexical Resource: $F(1, 64) = 0.92$, $p = .341$; Grammatical Range and Accuracy: $F(1, 64) = 0.48$, $p = .491$), which indicated homogeneity of regression slopes across all sub-construct models.

3.5.2 Qualitative Data Analysis

Braun and Clarke’s (2006) reflexive thematic analysis was used to analyze the reflective journals and FGD transcripts. All data were open-coded by the first author, yielding an initial codebook. A second coder analyzed a random sample of 30%, and a substantial level of inter-coder agreement was achieved, measured by Cohen’s κ (.79). Codes were grouped and iteratively refined into candidate themes, which were further developed and clarified through team discussions and then interpreted in relation to the quantitative results to develop explanatory accounts. Pseudonyms are used to refer to participants in the excerpts below.

The second coder was a doctoral-level applied linguist independent of the research team, trained on the codebook through two calibration sessions using non-study data before independently coding the 30% subsample. Trustworthiness was further supported through member checking, maintenance of a reflexive audit trail documenting analytic decisions, and deliberate attention to negative or disconfirming cases within the data set.

Ethical Approval

Ethical approval for this study was granted by the Research Ethics Committee of the College of Social Sciences and Humanities (approval no. CSSH/REC/2023-07), and written informed consent was obtained from all participants before data collection.

4 Results and Discussion

This section first presents the quantitative results addressing Research Questions 1 and 2, followed by the qualitative findings that explain these patterns in relation to Research Question 3, and closes with an integrated discussion of the findings in relation to the study's three objectives and the existing literature.

4.1 Quantitative Results

4.1.1 Group Equivalence at Pre-Test

An independent-samples t-test was used to test for pre-existing differences between groups on com-

posite writing scores, which confirmed that there was no difference ($t(66) = 0.84$, $p = .404$), thus supporting the validity of subsequent ANCOVA comparisons.

4.1.2 Effect of DST Intervention on Composite Writing Performance

Table 1 reveals that the experimental group achieved significantly higher composite post-test writing scores than the control group ($F(1, 65) = 18.43$, $p < .001$, $\eta^2 = .22$), after controlling for pre-test scores. The effect size suggests a large practical effect, with DST intervention explaining around 22% of the variance in the post-test writing scores.

Table 1: Adjusted Post-Test Means by Group (Composite Writing Score)

Group	n	Pre-Test M (SD)	Post-Test M (SD)	Adjusted M	SE
Experimental (DST)	34	4.61 (0.83)	6.34 (0.79)	6.31	0.13
Control	34	4.57 (0.88)	5.29 (0.91)	5.32	0.13

Note. Scores on 0–9 scale; $F(1, 65) = 18.43$, $p < .001$, $\eta^2 = .22$

4.1.3 Sub-Construct Analysis

An ANCOVA was performed for each of the four sub-constructs, with an assumption check performed for each. In all groups and a total of four sub-constructs, Shapiro–Wilk tests were not significant (all $p > .05$), suggesting no meaningful departure from normality. Levene's test indicated equal variances across groups for Task Achievement, $F(1, 66) = 0.02$, $p = .899$; Coherence & Cohesion, $F(1, 66) = 0.00$, $p = .986$; Lexical Resource, $F(1, 66) = 0.00$, $p = .974$; and Grammatical Range & Accuracy, $F(1, 66) = 0.23$, $p = .635$. The group $\times$ pre-test interaction term was not significant for any of the four sub-constructs: Task Achievement, $F(1, 64) =$

1.97, $p = .165$; Coherence & Cohesion, $F(1, 64) = 0.03$, $p = .869$; Lexical Resource, $F(1, 64) = 0.92$, $p = .341$; Grammatical Range & Accuracy, $F(1, 64) = 0.48$, $p = .491$, which indicated that regression slopes were homogeneous across all sub-construct models and that the standard ANCOVA approach reported in Table 2 was therefore appropriate.

As clearly shown in Table 2, separate ANCOVAs for each sub-construct revealed significant group differences for Coherence and Cohesion ($F(1, 65) = 22.17$, $p < .001$, $\eta^2 = .25$), Lexical Resource ($F(1, 65) = 15.88$, $p < .001$, $\eta^2 = .20$), and Task Achievement ($F(1, 65) = 12.49$, $p < .001$, $\eta^2 = .16$).

Table 2: ANCOVA Results by Writing Sub-Construct

Sub-Construct	F(1,65)	p	η^2	Significant?
Task Achievement (TA)	12.49	< .001	0.16	Yes
Coherence & Cohesion (CC)	22.17	< .001	0.25	Yes
Lexical Resource (LR)	15.88	< .001	0.2	Yes
Grammatical Range & Accuracy (GRA)	3.11	0.083	0.05	No

Note. Reported F-values are $F(1, 65)$; $\alpha = .05$.

Despite being in the expected direction, the group difference for Grammatical Range and Accuracy was not significant ($F(1, 65) = 3.11, p = .083, \eta^2 = .05$).

4.2 Qualitative Results: Explaining the 'Why'

To protect participant confidentiality, all names used in the excerpts below are pseudonyms (see Section 3.5.2); the 96 journal entries cited derive from the 12 Phase II participants described in Section 3.4.2. Thematic analysis of reflective journals and FGD transcripts resulted in four themes that explain how DST helps in developing academic writing. The following excerpts from participants support these themes.

4.2.1 Narrative Rehearsal as Cognitive Scaffolding

In each instance, participants reported that digital story planning was a form of cognitive preparation for constructing an academic argument. The need to sequence events causally, establish a point, and hold the audience's attention in a three-minute story appeared to internalize the structuring of reasoning, which was then transferred into the process of writing an essay.

Previously, I wrote and hoped for the ideas to merge and come together, but with the story, I had to ask, "Why is this happening?" What comes next? Then I began to write the same questions in my essay. (Selamawit, reflective journal, Week 5)

This finding is consistent with Bruner's (1990) account of narrative as a basic mode of thinking and provides evidence for the view that narrative is a developmental asset in expository and argumentative writing.

4.2.2 Multimodal Feedback Loops

Students tested the logic and precision of their ideas recursively across modes while composing multimodality, as they selected images, recorded voices, and timed transitions.

On recording my voice with the pictures and playing it back, I was able to hear when I

was not very clear in my explanation. I had to edit the script three times. I would never edit my essays so much. (Dawit, FGD 1)

This theme provides a rationale for the high achievement by students in Coherence and Cohesion, as they engaged in a multimodal process of externalizing the logic of the text, allowing them to identify and correct incoherence before it appeared in academic writing.

4.2.3 Affective Investment and Authorial Voice

The results indicated that participants were significantly more emotionally involved in their digital stories than in their traditional essays and that the presence of an authorial voice increased as they completed academic writing tasks.

In the essay, I had to pretend to be myself, but in the story, I was speaking as myself. After Cycle 3, I felt I could be myself in the essay as well, but in a more formal way. (Tigist, FGD 2)

This theme relates to Academic Literacies scholarship (Lea & Street, 2006) on identity development in academic writing and suggests that DST may serve as a space where students negotiate their personal and academic voices.

4.2.4 Collaborative Audience Awareness

Students felt this audience awareness was carried over from the peer-screening aspect of each DST cycle into writing their academic essays.

I knew that my classmates would watch my story, so I had to be clear, and while composing the essay, I always thought to myself, "Will the reader understand this?" That question was always in my head. (Biniyam, reflective journal, Week 7)

This theme, understanding that there is a real audience and that they have real information needs, may explain the relative gain in Task Achievement ($\eta^2 = .16$): understanding that a real audience exists and that they have real information needs will make students' writing more similar to the communicative purposes that the academic essay tasks are designed to simulate.

4.3 Discussions

The discussion below is organized around the study's three objectives and corresponding research questions, addressing the composite effect first, then the sub-construct pattern, and finally the explanatory mechanisms revealed by the qualitative strand.

Composite writing performance (Objective 1 / Research Question 1)

In line with the objectives in 1, the composite post-test scores in academic writing (Section 4.1.2) were found to be significantly higher for the experimental group, indicating that the structured DST intervention had a large practical effect when compared with the conventional teaching method. This result corroborates previous DST-writing research in other EFL contexts (Abdel-Hack & Helwa, 2014; Castañeda, 2013), thus generalizing their conclusions to another formal type of writing task (argumentative vs. narrative) and to another writing test setting (Ethiopia vs. other countries). The pattern also fits with more recent writing intervention literature that uses technology as a medium for intervention in similar tertiary EFL/ESL environments: Castillo-Cuesta *et al.* (2021) employed a quasi-experimental design ($n = 101$) in an EFL pre-service teacher training program in Ecuador, where five pre- and post-test groups were each given one of the four writing intervention methods examined, including DST, and the results showed that, in comparison with the other writing intervention groups, the DST group achieved significant gains in writing performance; and Xu (2021/2023) conducted a quasi-experimental study on the effectiveness of a digital multimodal composing intervention for Chinese university students ($n = 96$) in writing, and the findings revealed that the digital multimodal composing intervention resulted in significant improvement in five dimensions of L2 writing, including content and comprehensibility. On the one hand, the reviews in this section give further convergent support to the present research, as they concede that structured, technology-mediated composing interventions, including both DST and more general multimodal composing, generally improve student success with respect to writing, over and above conventional instruction, in a variety of tertiary EFL settings.

Sub-construct pattern (Objective 2 / Research Question 2)

The large effect size for coherence is especially remarkable because coherence, or the logical connection between ideas within texts, is often considered one of the most persistent problems in the writing of EFL academic writers (Hyland, 2019). The qualitative data indicate that the multimodal language of DST externalized coherence demands in a manner not found in the written drafts, which seemed to be transferred to the process of essay composition, a process that is also seen as an example of self-regulated learning (Zimmerman, 2000). The results are expected within the theoretical framework, as the difference between the groups is not statistically significant in terms of grammatical range and accuracy. Based on scaffolding theory, the competencies that will respond best to a particular scaffold will be the ones directly stimulated by the scaffolding activity. When considering grammatical accuracy gains, gains are expected; however, given the emphasis on communicative coherence and purposeful use of language, such gains would not be expected to be statistically significant within the time frame of the study.

Explanatory mechanisms (Objective 3 / Research Question 3)

The qualitative outcome of affective investment as a scaffold to the development of writing enriches the cognitive view of scaffolding for writing. The participants' accounts of the emergence of an authorial voice in DST confirm Ivanič's (1998) idea of the "writer identity" as an ongoing construction that is negotiated within the specific context. This is consistent with recent empirical studies of L2 writer identity that trace the evolution of L2 writer identity over multiple drafts of a thesis (Zhao & Liu, 2021) or multiple trials of a second language (Zhang & Wang, 2023), showing that L2 writer identity was not formed at once but was rather constructed — that is, developed through revision and feedback from advisors — with L2 writer identity becoming more independent as writers gained more experience. It is important to note that some students explicitly identified this discovery of voice in the transition from Cycle 3 (argumentative narrative) to Cycle 4 (academic abstract), implying that the

scaffolding sequence used to strategically shorten the gap between the two cycles was effective not only in transferring skills but also in structuring the transition of identity.

Taken together, the quantitative and qualitative strands converge: the composite and sub-construct gains (Objectives 1–2) are explained by the four mechanisms identified in Section 4.2 (Objective 3), supporting the conceptualization, advanced in Section 2.1, of the four-cycle DST sequence as a macro-scaffold that incrementally withdraws support as students move from personal narrative toward formal academic register.

5 Conclusions and Recommendations

The present sequential explanatory mixed-methods study shows that a structured digital storytelling intervention is effective in enhancing EFL undergraduates' academic writing in the following aspects: coherence and cohesion, lexical resources, and task achievement, with particular effects on the first two aspects. The qualitative strand, in addition to documenting the effect, uncovers four explanatory mechanisms that illuminate how DST can serve as a transitional scaffold: narrative rehearsal, multi-modal feedback loops, affective investment, and awareness of the audience.

5.1 Practical Implications

The results have a number of implications for practice. Instructors of academic writing in EFL tertiary settings might be well advised to embed structured DST cycles in the early stages of academic writing classes, developmentally moving from personal writing to disciplinary argumentation. The present intervention, in which the Cycle 3–4 transition is from argumentative narrative to formal academic abstract, seems to be a pedagogically rich transition moment and deserving of further study as an independent pedagogical technique.

The study indicates that, for curriculum designers, DST could be most useful as an adjunct to the conventional approach to teaching and learning writing, rather than as its substitute, in order to engage the affective, cognitive, and social aspects of writing that conventional approaches may not

foster. Process writing was still carried out in the control group, providing an intensive comparison condition, and the results suggest that the two teaching methods together may be more effective than either method alone.

5.2 Limitations and Directions for Future Research

The study's limitations are as follows: assignment to condition occurred at the level of intact classes rather than individuals, which introduces a risk of cluster confounds despite the pre-test equivalence check (Section 4.1.1); the achieved sample size was below the benchmark suggested by an a priori power analysis (Section 3.2); there is a single-institution setting; the intervention lasted 8 weeks; and the Phase II participants were self-selected and may not fully represent the diversity of the experimental group's experience. Further studies are needed on writing scaffolds based on DST in different disciplines, on a longer time frame, and in culturally diverse EFL environments, ideally using individual-level randomization and larger, adequately powered samples. Longitudinal studies that follow writers over more than one course would be especially useful to assess whether the scaffolding shown here continues to strengthen and deepen.

Acknowledgements

The author wishes to thank the students who generously participated in this study and the EAP instructors who facilitated data collection.

Conflict of Interest

The author declares no conflict of interest with respect to the research, authorship, or publication of this article.

References

- Abdel-Hack, E. M., & Helwa, H. S. A. A. (2014). Using digital storytelling and weblogs instruction to enhance EFL narrative writing and critical thinking skills among EFL majors at the Faculty of Education. *Educational Research*, 5(1), 8–41.

- Ajabshir, Z. F. (2024). Empowering EFL writing through digital storytelling: A quasi-experimental assessment of CALF measures and multidimensional engagement. *Acta Psychologica*, 250, Article 104564. <https://doi.org/10.1016/j.actpsy.2024.104564>
- Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. Routledge.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- British Council. (2022). *IELTS writing band descriptors (Task 2)*. British Council/IDP/Cambridge Assessment English.
- Bruner, J. (1990). *Acts of meaning*. Harvard University Press.
- Castañeda, M. E. (2013). ‘I am proud that I did it and it’s a piece of me’: Digital storytelling in the foreign language classroom. *CALICO Journal*, 30(1), 44–62. <https://doi.org/10.11139/cj.30.1.44-62>
- Castillo-Cuesta, L. M., Quinonez-Beltran, A., Cabrera-Solano, P., Ochoa-Cueva, C., & Gonzalez-Torres, P. (2021). Using digital storytelling as a strategy for enhancing EFL writing skills. *International Journal of Emerging Technologies in Learning (iJET)*, 16(13), 142–156. <https://doi.org/10.3991/ijet.v16i13.22187>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.
- Dyson, A. H., & Genishi, C. (1994). *The need for story: Cultural diversity in the classroom and community*. National Council of Teachers of English.
- Fantaye, D. B. (2020). Academic writing problems of Ethiopian university students: The case of Addis Ababa University. *Ethiopian Journal of Education and Sciences*, 15(2), 1–18.
- Gibbons, P. (2015). *Scaffolding language, scaffolding learning: Teaching English language learners in the mainstream classroom* (2nd ed.). Heinemann.
- Hammond, J. (2020). Scaffolding and language: Challenges and debates. In K. Scrimgeour (Ed.), *Language and learning in primary school* (pp. 45–67). Routledge.
- Hammond, J., & Gibbons, P. (2005). Putting scaffolding to work: The contribution of scaffolding in articulating ESL education. *Prospect*, 20(1), 6–30.
- Hyland, K. (2019). *Second language writing* (2nd ed.). Cambridge University Press.
- Ivanič, R. (1998). *Writing and identity: The discursive construction of identity in academic writing*. John Benjamins.
- Lambert, J. (2018). *Digital storytelling: Capturing lives, creating community* (5th ed.). Routledge.
- Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, 23(2), 157–172. <https://doi.org/10.1080/03075079812331380364>
- Lea, M. R., & Street, B. V. (2006). The ‘academic literacies’ model: Theory and applications. *Theory Into Practice*, 45(4), 368–377. https://doi.org/10.1207/s15430421tip4504_11
- Lillis, T., & Curry, M. J. (2010). *Academic writing in a global context: The politics and practices of publishing in English*. Routledge.
- Mulugeta, A., & Getnet, B. (2022). Challenges of academic writing in Ethiopian universities: A systematic review. *Journal of Language Teaching and Research*, 13(4), 723–734. <https://doi.org/10.17507/jltr.1304.03>
- Nguyen, L. T. (2021). Digital storytelling in second language learning: A systematic review. *CALL-EJ*, 22(2), 88–114.
- Ohler, J. B. (2013). *Digital storytelling in the classroom: New media pathways to literacy, learning, and creativity* (2nd ed.). Corwin Press.

- Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4th ed.). SAGE.
- Robin, B. R. (2016). The power of digital storytelling to support teaching and learning. *Digital Education Review*, 30, 17–29. <https://doi.org/10.1344/der.2016.30.17-29>
- Sylvester, R., & Greenidge, W. (2009). Digital storytelling: Extending the potential for struggling writers. *The Reading Teacher*, 63(4), 284–295. <https://doi.org/10.1598/RT.63.4.3>
- Vinogradova, P., Linville, H. A., & Bickel, B. (2011). ‘Listen to my story, and you will know me’: Digital stories as student-centred collaborative projects. *TESOL Journal*, 2(2), 173–202. <https://doi.org/10.5054/tj.2011.250213>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Xu, Y. (2023). Investigating the effects of digital multimodal composing on Chinese EFL learners’ writing performance: A quasi-experimental study. *Computer Assisted Language Learning*, 36(4), 785–805. (Published online 2021.) <https://doi.org/10.1080/09588221.2021.1945635>
- Yimam, B. (2018). Features of academic writing by Ethiopian university students: Problems and pedagogical implications. *Ethiopian Journal of Languages and Literature*, 7(1), 1–27.
- Zhao, J., & Liu, Y. (2021). A developmental view of authorial voice construction in master’s thesis: A case study of two novice L2 writers. *SAGE Open*, 11(4), 1–15. <https://doi.org/10.1177/21582440211054483>
- Zhang, [initial], & Wang, [initial]. (2023). To be like a "scholar": A study on the construction of authorial identity of Chinese EFL learners in academic writing: An intertextuality perspective. *Frontiers in Psychology*, 14, Article 1297557. <https://doi.org/10.3389/fpsyg.2023.1297557>
- Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25(1), 82–91. <https://doi.org/10.1006/ceps.1999.1016>
- Zinsser, W. (2013). *Writing to learn*. Harper Perennial.