Article

The Effects of Generative AI Platforms on Undergraduates' Narrative Intelligence and Writing Self-Efficacy

Nikolaos Pellas 🄳

Department of Primary Education, University of Western Macedonia, 53100 Florina, Greece; aff00192@uowm.gr

Abstract: Digital storytelling and generative artificial intelligence (AI) platforms have emerged as transformative tools that empower individuals to write with confidence and share their stories effectively. However, a research gap exists in understanding the effects of using such web-based platforms on narrative intelligence and writing self-efficacy. This study aims to investigate whether digital story creation tasks on web-based platforms can influence the narrative intelligence and writing self-efficacy of undergraduate students. A pretest–posttest comparison study between two groups was conducted with sixty-four undergraduate students ($n = 64$), majoring in Primary Education. More specifically, it compares the effects of the most well-known conventional platforms, such as Storybird, Storyjumper, and ZooBurst (control condition), and generative AI platforms, such as Sudowrite, Jasper, and Shortly AI (experimental condition) on undergraduate students, with an equal distribution in each group. The findings indicate that the utilization of generative AI platforms in the context of story creation tasks can substantially enhance both narrative intelligence scores and writing self-efficacy when compared to conventional platforms. Nonetheless, there was no significant difference in the creative identity factor. Generative AI platforms have promising implications for supporting undergraduates' narrative intelligence and writing self-efficacy in fostering their story creation design and development.

Keywords: artificial intelligence; digital platforms; narrative intelligence; writing self-efficacy; storytelling

1. Introduction

Digital storytelling creation, a contemporary evolution of a timeless art, requires the utilization of digital tools and multimedia formats to breathe life into narratives. This empowers creators to seamlessly weave together a wide array of multimedia elements, resulting in compelling and immersive stories [1]. It can also take various forms and be delivered through different multimedia features using a wide range of digital platforms, including written or spoken words, images, videos, audio, and interactive elements, supporting students to develop writing skills, creativity, and digital literacy [2].

A diverse array of digital platforms exists for creators to craft engaging and immersive stories, enabling them to harness multimedia elements to bring their narratives to life. On the one side, web-based storytelling platforms can be a valuable resource for teaching students how to create descriptive audiovisual storytelling. Storybird (https://storybird.com, (accessed on 10 October 2023), Storyjumper (https://www.storyjumper.com, (accessed on 10 October 2023), and Mixbook (https://www.mixbook.com, (accessed on 8 October 2023) are among the most prominent and common-in-use platforms intentionally crafted to streamline the process of story creation and sharing [3]. These platforms boast an extensive selection of templates and user-friendly tools which enable individuals to craft and embellish their unique narratives [4]. Platforms for story creation are specially tailored to cater to the needs of educators, young students, and budding authors, providing them with a conducive environment to unleash their creative potential and expand their knowledge across various disciplines. For example, Storyjumper was recommended by Ispir and Yıldız [5] as an innovative method of teaching digital storytelling and writing to students. Utilizing these platforms offers students several

advantages, including a significant improvement in their scores for descriptive text writing [6]. Furthermore, Fitriyani et al. [7] advocated that both teachers and students had a favorable perception of Storyjumper's usage in the classroom.

On the other hand, machine learning (ML) and natural language processing (NLP) technologies, extensively employed within artificial intelligence (AI) technologies, have now become indispensable components of digital platforms, influencing numerous facets of humans' everyday existence [8]. Moreover, large language models (LLM) powered by AI are meticulously designed to produce text that closely resembles human writing by analyzing extensive datasets gathered from the Internet [9]. The swift progression of AI and ML innovations has not only revolutionized human–computer interaction using digital platforms, but has also sparked essential inquiries concerning the societal consequences of these technologies [10]. AI story generators represent revolutionary generative AI platforms engineered to liberate authors from the clutches of writer's block and elevate the quality of their narratives, effortlessly conjuring up a treasure trove of ideas, characters, and plotlines [11]. Generative AI is considered a subset of AI technology that can produce original content by learning from data and using complex algorithms and neural networks to create human-like text, images, and music [8,10]. As cutting-edge technologies, AI platforms boast a user-friendly interface that ensures accessibility for writers at all skill levels to enhance the writing process, empowering users to craft content that is both captivating and uniquely their own to weave narratives that transcend their creative boundaries and enrich their literacy creations [12]. Sudowrite (https://www.sudowrite.com, (accessed on 20 October 2023), Jasper (https://www.jasper.ai, (accessed on 19 October 2023), and Shortly AI (https://www.shortlyai.com, (accessed on 19 October 2023) are some of the latest platforms that offer meticulous writing refinement assistance by suggesting alternative phrasing, synonyms, and sentence structures. These tools suggest alternative phrasing, synonyms, and sentence structures to enhance the overall quality of the text. By identifying and proposing improvements for clichéd expressions and overused phrases, they assist writers in revitalizing their work with unique and original content [13]. Lee et al. [14] revealed that AI-generated content was notably more effective and preferable over traditional English-as-a-foreign-language reading instruction. Im et al. [15] admitted the impact of generative AI story relay that emerged as a powerful tool in promoting developers' comprehension of and regard for user perspectives, while also fostering users' critical thinking regarding the societal implications of AI. Lastly, generative AI platforms for video content creation were recommended by Pellas [16] as innovative tools that can benefit the digital storytelling and writing abilities of undergraduates regardless of their socio-cognitive backgrounds.

Digital storytelling combines the art of storytelling with the dynamic capabilities of multimedia elements provided by digital platforms. Educators and instructors can use generative AI platforms or web-based platforms to create stories that transcend traditional text-based narratives. These stories can incorporate images, videos, audio, and interactive elements to create engaging and immersive experiences [17]. to establish profound bonds with their audience, eliciting emotions and understanding on various sensory planes. In this dynamic realm of storytelling, the significance of writing self-efficacy cannot be overstated. This concept denotes an individual's confidence in their ability to adeptly strategize, create, and refine written compositions [18]. It stands as a fundamental psychological element that significantly influences the writing process and overall writing proficiency. High levels of writing self-efficacy empower individuals to approach digital storytelling endeavors with assurance and determination, enhancing the quality of their narratives [19]. Nurtured through constructive experiences, feedback, and supportive learning environments, writing self-efficacy contributes to the development of proficient writers capable of conveying their ideas persuasively and effectively across diverse digital contexts, from educational platforms to professional communication channels [4].

Through the fusion of technology and narrative, digital storytelling not only entertains but also educates, informs, and elicits emotions, thereby reshaping how stories are both

created and consumed in the digital age. Narrative intelligence, often linked with human cognition, refers to the capacity to understand, construct, and navigate narratives as another potential factor affecting writing self-efficacy. Randall [20] defines narrative intelligence as the capacity to comprehend and craft narrative frameworks. While emotional and verbal intelligence has been extensively examined in the context of language acquisition over the last decade, narrative intelligence has not received as much scrutiny [21]. Humans are inherently wired for storytelling, and narrative intelligence is what enables people to comprehend complex stories, discern patterns, and make sense of the world around us. It involves identifying narrative structures, recognizing characters, following plotlines, and extracting meaning from stories [22,23].

In recent years, the concept of narrative intelligence has extended to storytelling creation platforms, where digital-oriented and generative AI platform features and elements exist to not only process text and data but also to understand and generate narratives in a coherent and contextually relevant manner. More specifically, a generative AI platform's narrative intelligence has the potential to find applications in content creation, chatbots, virtual assistants, and even the storytelling aspect of video games. This contributes to creating digital experiences that are more engaging and immersive [1,8]. In contrast, traditional storytelling relies solely on human creators for content generation. The increasing significance of technology integration has made it imperative to judiciously and effectively incorporate contemporary technologies into various educational domains. Furthermore, the incorporation of technology into distinct academic domains should be harmonized with the distinctive teaching and learning methodologies intrinsic to those fields [24]. This integration significantly impacts individuals' writing self-efficacy and narrative intelligence. It is anticipated that this kind of modern technology will have a substantial impact on literacy education, given its pivotal role in the realm of innovative instruction within a substantial range of disciplines [25].

However, the use of generative AI platforms to support learning and foster creativity in educational contexts is still in its early stages of research. Questions persist regarding the effectiveness, validity, and safety of these AI tools, as indicated in references [7,26]. This is particularly crucial in the ever-changing educational landscape, where comprehending how learners perceive their interactions with AI and the strategies they employ to collaborate with it is vital for the successful integration of these systems. While a growing body of literature [13–15] has investigated the relationship between writing self-efficacy, narrative intelligence, and storytelling in the context of digital platforms, a notable research gap exists in understanding whether the integration of emerging technologies, such as generative AI platforms, influences individuals' self-perceived writing abilities and narrative competence using digital storytelling platforms. This research gap becomes increasingly relevant as generative AI tools for narrative creation become more prevalent on digital platforms. While several studies [8–10,15] have examined the potential benefits and drawbacks of generative AI writing, there is a need for comprehensive investigations that consider the nuanced interplay between human writers' self-efficacy, their narrative intelligence, and the extent to which AI tools influence their storytelling experiences.

Based on the above, two research questions can be formulated as follows:

RQ1. Does creating digital stories using generative AI platforms have a significant impact on the narrative intelligence of undergraduate students?

RQ2. Does creating digital stories with generative AI platforms have a significant effect on the writing self-efficacy of undergraduate students?

This study aims to explore the consequences of digital story creation tasks on the levels of narrative intelligence and writing self-efficacy among undergraduates majoring in Primary Education using several web-based storytelling platforms. Addressing this research gap is essential for a more holistic understanding of whether the evolving landscape of digital storytelling and generative AI platforms influences individuals' self-efficacy in writing and their narrative intelligence. Furthermore, insights from such research can

inform the design of generative AI tools and platforms that enhance rather than hinder individuals' storytelling capabilities and confidence in the digital age.

2. Materials and Methods

2.1. Research Design

This study used a pretest–posttest comparison design with two groups of participants, a well-known and reliable design in educational research to examine the effects of a treatment between two groups [27]. Pretests and posttests were used to measure the dependent variable (i.e., digital story creation) before and after the treatment. In the pretest phase, before the intervention, participants' baseline narrative intelligence and writing self-efficacy were assessed. This study aimed to evaluate the impact of the intervention (i.e., the use of generative AI platforms) on the participants' narrative intelligence and writing self-efficacy. Any observed differences in the posttest scores would be attributed to the intervention itself, rather than preexisting disparities between the groups. In other words, the results would suggest if students' engagement with generative AI platforms can lead to significant improvements in these measures compared to the control group. which used conventional platforms.

While the independent variable of the research was the applied instructional method when students used different platforms, the dependent variables were measured using narrative intelligence and writing self-efficacy scales. The method of research is shown in Table 1.

Table 1. The proposed research design.

Groups	Pretest	Treatment	Posttest
G1 (NR)	O_1		O_2
G2 (NR)	O_3	X	O_4

G1: control group—story creation using a conventional platform. G2: experimental group—digital story creation using generative AI. O_1 and O_3 are the pretests for G1 and G2, respectively. O_2 and O_4 are the posttests for G1 and G2, respectively. X: digital story creation.

In this research design, NR (non-randomization) signifies the deliberate decision not to randomize participants [28,29]. It is of great importance to note that although only the experimental group received the treatment, both the experimental and control groups completed both the pre-and posttests. This non-random assignment of participants to the two groups was a deliberate choice made to prevent potential bias in the study's results [30,31]. The main researcher wanted to avoid having an experimental group consisting only of experienced users because it would have made it difficult to determine whether an improvement in narrative intelligence and writing self-efficacy was due to the use of generative AI platforms or simply because the participants were already experienced users.

To reduce the potential influence of novelty, the researcher ensured that all participants had some experience with using digital platforms correctly. The experimental group used generative AI digital platforms for story creation, while the control group used more traditional platforms [28,32]. This allowed the researcher to compare the effects of generative AI platforms to the effects of traditional platforms while controlling for the potential influence of novelty.

2.2. Sampling

The research utilized a convenience sampling method [27]. During the spring semester of the 2022–2023 academic year, a total of sixty-four (*n* = 64) students, majoring in Primary Education, were selected as participants. All were Greek university students specializing in educational science courses, with a specific emphasis on primary school education. The courses encompassed a diverse array of subjects, including computer science, as well as visual and auditory instructional design in various STE(A)M disciplines.

The experimental group consisted of thirty-two participants (*n* = 32) and the control group also comprised thirty-two participants (*n* = 32). This is a relatively small sample size,

but it is sufficient to test the feasibility of a larger study. The results of a pilot study can be used to improve the design of a subsequent larger study [32]. By carefully controlling the variables in the study, the researcher could be more confident that the results are due to the use of generative AI platforms and not to other factors. Within this context, the participants utilized a wide range of generative AI and "conventional" platforms to create a digital story aligned with the Greek literacy curriculum's learning objectives for elementary school students.

The criteria for including these specific students were based on their enrollment in the Department of Primary Education at a university in Greece during the spring semester of the 2022–2023 academic year. These students were chosen because they represented a feasible and accessible research group within the given time frame and resources. The 64 students selected for this study were drawn from a Department of Primary Education at a Greek university, and they participated in the study voluntarily. However, it is important to note that this sample may not be fully representative of the entire department's student population. The sample size in this study was relatively small, and its representativeness may be limited. There are a larger number of students in the Department of Primary Education, but for practical reasons, a smaller sample was chosen to conduct the current study. The exact number of students in the Department of Primary Education during the specified academic year was not formally provided to the instructors. However, this information could be obtained from the department's records or administration for a more comprehensive understanding of the sample's size about the department's overall student population. While a sample size of 64 participants may be considered relatively small, it was selected as a "pilot" to test the feasibility of a larger-scale investigation [27]. The results can inform the design of subsequent, more extensive research. Careful control of variables in the study assured the researcher that the results are indicative of the effects of generative AI platforms on narrative intelligence and writing self-efficacy. The decision to use non-random assignment in this study was intentional and aimed at preventing potential bias in the results [28]. By including both the experimental and control groups in the pre-and posttests, any observed differences could be more confidently attributed to the use of generative AI platforms rather than prior experience. Non-random assignment allowed the researchers to compare two groups with different levels of familiarity with digital storytelling platforms, contributing to a more robust assessment of the impact of generative AI-supported instructional methods.

The gender distribution of the participants is detailed in Table 2, and their distribution based on their general weighted grade averages is presented in Table 3.

Table 2. Distribution of participant groups by gender.

Gender	Control Group		Experimental Group		Total	
	n	%	*n*	%	*n*	%
Female	20	60.4	23	69.8	43	64.1
Male	12	36.6	9	31.2	21	35.9
Total	32	100	32	100	64	100

Table 3. Results of chi-square analysis.

Group	Gender				Total		X²	df	*p*
	Female		Male						
	n	%	*n*	%	*n*	%			
Control	19	29.7	13	20.3	32	50	0.58	1	0.455
Experimental	22	34.4	10	15.6	32	50			
Total	41	64.1	23	35.9	64	100			

When Table 2 is examined, it can be seen that the percentages of female participants in the control group (60.4%) and the experimental group (68.8%) were close. Likewise, the percentages

of male participants in the control group (40.6%) and the experimental group (36.6%) were close. The majority of the total participants were women (69.8%). A chi-square analysis was conducted to examine whether the participant groups differed according to gender.

When Table 3 is examined, it can be concluded that the participant groups did not differ according to gender. A chi-square analysis was conducted to examine whether the participant groups were dependent on gender. The chi-square analysis results in Table 3 present the relationship between gender and group membership (control and experimental). The table reveals that both groups consisted of an equal number of participants. The distribution of gender within these groups shows that there were slightly more women in the experimental group (34.4%) compared to the control group (29.7%), while the latter group had a slightly higher proportion of men (20.3%) compared to the experimental group (15.6%). However, a chi-square test with a statistic (x^2) of 0.58 and a $p = 0.455$ indicated that there was no statistically significant association between gender and group membership in this sample, suggesting that gender distribution in both groups was not significantly different from what would be expected by chance.

An independent sample *t*-test was carried out to investigate if there were variations in the pretest scores related to writing self-efficacy between the undergraduate students in the experimental and control groups. In Table 4, the control group consists of 32 participants with a mean (M) SAWSES score of 3.66 and a standard deviation (SD) of 0.53. The experimental group, with the same number of participants, had a slightly higher mean SAWSES score of 3.77 and a lower standard deviation of 0.44. A two-sample *t*-test was conducted to compare the means of the two groups, resulting in a t-statistic of -1.88 and a *p*-value of 0.11. The t-value indicates that the control group's pretest scores are lower than the experimental group, although this difference is not statistically significant at the 0.05 alpha level ($p > 0.05$). Consequently, there is no strong evidence to suggest a significant difference in pretest SAWSES scores between the two groups.

Table 4. Comparison of experimental and control group SAWSES (pretest scores).

Group	*n*	M	SD	df	t	*p*
Control Group	32	3.66	0.53	62	−1.88	0.11
Experimental Group	32	3.77	0.44			

The researcher also took strict ethical considerations into account when conducting this study. First, to ensure a diverse participant group, purposive sampling was employed, involving the selection of individuals with varying experiences across different academic disciplines and a range of digital proficiency levels. This decision was made to avoid any potential disparities in digital skills among the participants, which could have affected the results of the study. The researcher also took steps to reduce potential biases and ensure the internal validity of the study by standardizing the participants' demographic characteristics and ensuring that all of them had substantial experience with AI-generated story content creation. Second, informed consent was obtained from all participants, ensuring confidentiality and anonymity and safeguarding the well-being and privacy of the participants. Voluntary participation was the only method of involvement, and all participants were required to provide informed consent before data collection began. The researcher also explained the study's objectives to the participants and outlined the potential consequences of using assessment platforms, the collection and handling of data following the General Data Protection Regulation (GDPR), and each participant's right to withdraw from the study at any time without facing adverse consequences.

2.3. Measures

2.3.1. Narrative Intelligence Scale

To assess the narrative intelligence of students, the researchers utilized the Narrative Intelligence Scale (NIS), which was developed and validated by Pishghadam et al. [23]

following the model proposed by Randall [20]. The NIS demonstrated high item reliability, at 0.99, and strong person reliability, at 0.87, with each item being rated on a scale of 1 to 5. The narrative intelligence of undergraduate students was evaluated by having them engage in two narrative tasks [29]. For the first task, they were instructed to discuss a strip story. In the second phase, the students were asked to narrate their experiences to create a story based on Greek book selections and to assist younger students in understanding some of its components. These narratives were subsequently recorded, transcribed, and assessed based on their narrative intelligence. In this research, the overall reliability of the questionnaire, as estimated through Cronbach's alpha, was found to be 0.81. The five factors of narrative intelligence along with their corresponding descriptions are as follows:

1. Emplotment involves the process of organizing events, especially those that occur less frequently or are of high significance, into a coherent and meaningful order. It helps create a narrative structure that makes sense to the audience;
2. Characterization focuses on the ability to vividly depict and create a mental image of the thoughts, emotions, and personalities of the participants or characters involved in a narrative. It adds depth and relatability to the story;
3. Narration is the skill of effectively conveying information and engaging in a dialogue with others about the events taking place in a narrative. It involves presenting events logically and making assumptions that allow for meaningful communication;
4. Generation entails the process of arranging events within a narrative in a way that makes them predictable and coherent. It involves creating a structured and understandable sequence of events;
5. Thematization refers to the ability to recognize and be aware of recurring patterns or themes within specific events. It involves identifying common elements or structures that help shape the overall meaning of a narrative.

In the present study, reliability scores above 0.70 or higher for test scores were considered acceptable for all questionnaires [30]. The reliability coefficients for these factors were found to be as follows: emplotment $\alpha = 0.86$, characterization $\alpha = 0.85$, generation $\alpha = 0.82$, narration a = 0.88, and thematization $\alpha = 0.87$. For the scale overall, the reliability coefficient was $\alpha = 0.84$. These results confirmed that the scale used in this study demonstrates reliability, with an alpha coefficient exceeding 0.70.

2.3.2. Situated Academic Writing Self-Efficacy Scale

In this study, the "situated academic writing self-efficacy scale (SAWSES)" developed by Mitchell et al. [31] was also employed. The researchers collected data from a sample of 543 undergraduate students during the development phase of the scale to validate this instrument. They utilized various statistical techniques, including exploratory factor analysis for component analysis, item–total correlation to assess reliability, *t*-tests to measure differences between groups in the top and bottom 27%, Spearman–Brown two half-test correlations, and Cronbach's alpha for internal consistency coefficient calculations.

This scale comprises 17 items and is divided into three factors: writing essentials (3 items), relational reflective writing (8 items), and creative identity (6 items). The reliability coefficients for these factors were found to be as follows: writing essentials $\alpha = 0.89$, relational reflective writing $\alpha = 0.87$, and creative identity $\alpha = 0.82$. For the scale overall, the reliability coefficient was $\alpha = 0.83$. The reliability results for this research confirmed that the scale used in the study exhibits strong reliability ($\alpha > 0.70$, see recommendations by Cortina [32]).

2.4. Procedure

In the course of this study's procedure, students in the experimental group were tasked with crafting generative AI-supported digital stories, while undergraduate students in the control group were instructed to create their own stories using "conventional" platforms. The experimental group received training from the researcher on crafting digital stories and digital storytelling for a potential school tool in literacy education utilizing Storybird,

Storyjumper, ZooBurst (control condition), and AI-generated platforms, such as Sudowrite, Jasper, and Shortly AI (experimental condition). Initially, the instructor (main researcher) provided a foundational understanding of the curriculum's learning objectives, and input was sought throughout the process from literacy education instructors. To investigate the impact of various platforms on students' learning outcomes and to explore potential connections between narrative intelligence and writing self-efficacy, an experimental study was conducted in a controlled laboratory environment and after hours in online courses.

The digital story creation process followed the 5 stages outlined by Cennamo et al. [29]. Leveraging the power of generative AI platforms, users were guided through the process of crafting visual stories, fostering the development of their narrative skills. These platforms helped users to create visually engaging stories that told compelling narratives. By providing users with a variety of tools and resources, these platforms helped users to develop their storytelling skills and create stories. To address the potential shortcomings of existing AI art tools, the researcher introduced sticker-based interfaces in which users can select stickers to personalize their stories. Once a user selected a sticker, the system responded by sending carefully curated prompts in the form of captions for the stickers. These prompts included detailed descriptions of the images, along with highlighted keywords and styles, such as digital arts, as a precautionary measure against generating inappropriate content. Additionally, some key points highlighting the significance of the participants' story creation are as follows [12,14–16]:

- Promoting inclusivity and equity: This study recognizes that digital storytelling has the potential to democratize content creation and media production. By examining how these technologies are used by undergraduate students from diverse sociodemographic backgrounds, this study contributes to understanding whether story creation tasks can help bridge gaps in inclusivity and equity in the digital media landscape;
- Impact of storytelling training courses: This study underscores the importance of training courses in shaping students' attitudes towards story creation technologies. This finding has implications for educational institutions and policymakers who may consider integrating Greek curricula to prepare students for the evolving digital media landscape;
- Quality and reliability: By assessing students' contentment with the reliability of creative storytelling technologies, this study highlights the importance of ensuring that story content meets certain quality standards based on lessons taught from the Greek curriculum. This is significant for media producers and policymakers maintaining the integrity of multimedia platforms;
- Policy considerations: This research emphasizes the need for considerations and policy guidance in instructional contexts to create a fair and equitable digital media environment. This is particularly relevant in an era where story creation technologies are becoming increasingly prevalent in media production.

All web-based questionnaires (demographics, narrative intelligence, and SAWSES) used to collect data from participants were self-reported, delivered to participants via email, and could be completed in no more than 50 min. This was designed to avoid categorizing participants as novice or expert users so that all participants had an equal opportunity to answer the questions and their responses were not biased by their level of experience. The researcher created a weekly lesson plan that guided the entire process. The SAWSES was administered as both a pretest and posttest to assess undergraduate students in both the experimental and control groups. Additionally, the narrative intelligence questionnaire was administered as a posttest after the study (Figure 1).

2.5. Data Analysis

The analysis of the research data involved several statistical techniques. Descriptive statistics were used to analyze and summarize the descriptive findings. Cronbach's alpha (α) was employed to assess the reliability of the scales. Additionally, the Kolmogorov–Smirnov and Shapiro–Wilk tests were applied to assess the distribution of the data. An independent sample *t*-test was used to investigate whether there were significant differences

in writing self-efficacy between groups. The first research question was addressed using an independent sample *t*-test, while the second research question was answered using a two-way ANOVA. The IBM® SPSS Statistics (version 33) program was employed to conduct a comprehensive and meticulous analysis of the data and scrutinize every detail utilizing a variety of statistical methods.
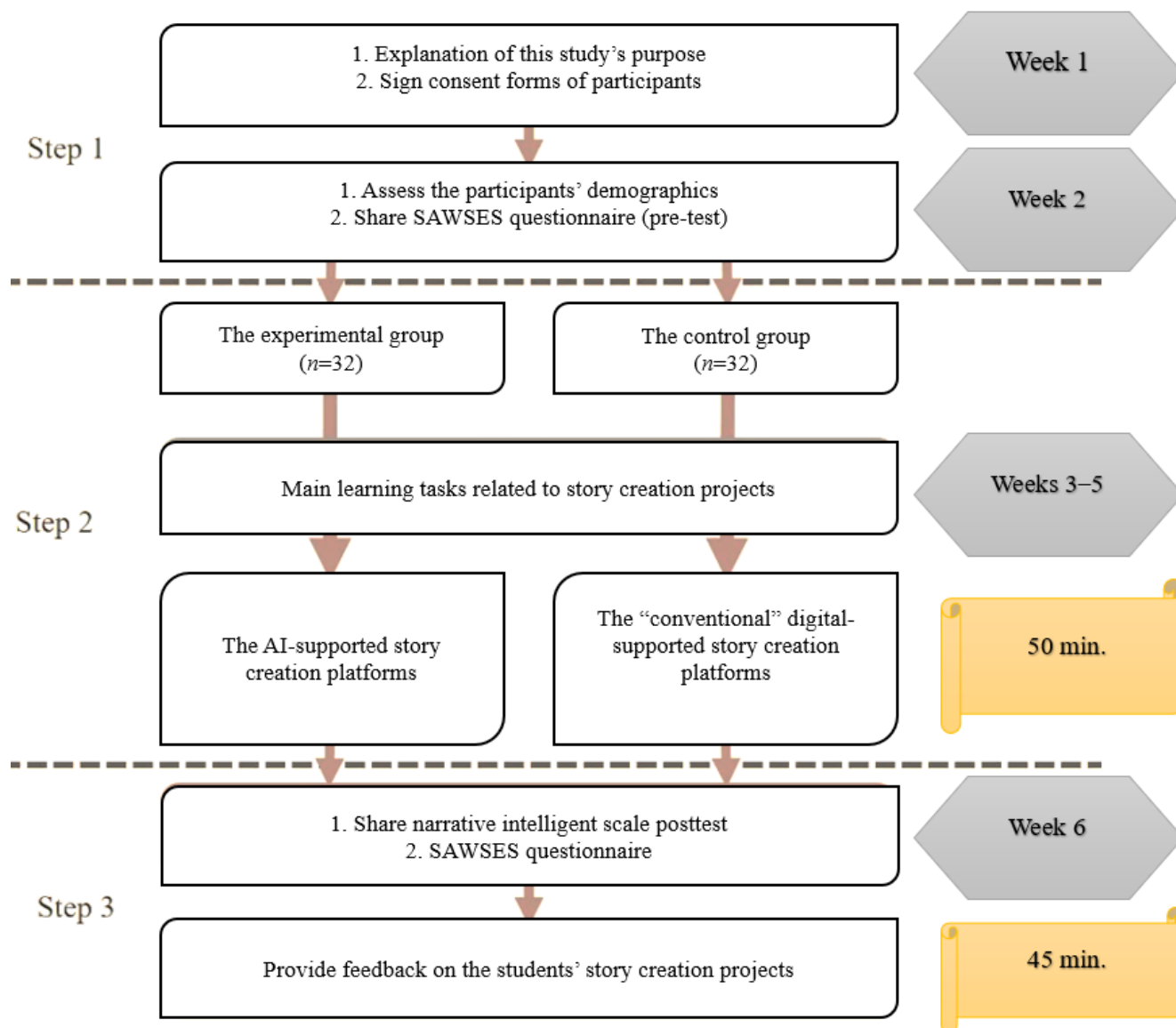

Figure 1. The main procedure.

The Kolmogorov–Smirnov and Shapiro–Wilk tests were carried out to assess whether the data exhibited a normal distribution, following the guidelines of Garth [33]. Table 5 provides the distribution parameters of the research data. If the *p*-values obtained from the Kolmogorov–Smirnov and Shapiro–Wilk tests exceed 0.05, it indicates that the data follow a normal distribution [33]. Upon examining the scales in Table 5, it was observed that the data exhibited a normal distribution since *p*-values were greater than 0.05 for both groups.

The back-translation method was utilized for all questionnaires in Greek to ensure that the translation was accurate by having it translated into the target language and then translated back into the original language. This helps to identify any errors or inconsistencies in the translation [34]. All 5-point Likert scales asked respondents to rate their level of agreement or disagreement with a statement on a scale of 1 to 5. The higher the score, the more agreement with the statement. There are a set of statistical analyses that can be used

to analyze data from Likert scales. These analyses can be used to identify patterns in the data, such as which statements are most or least agreed with, and to compare the responses of different groups of respondents.

Table 5. Distribution parameters of research data.

Test	Group	Kolmogorov–Smirnov		Shapiro–Wilk	
		N	*p*	N	*p*
Writing Self-Efficacy Scale pretest	Control	32	0.23	32	0.71
	Experimental	32	0.23	32	0.17
Writing Self-Efficacy Scale posttest	Control	32	0.23	32	0.76
	Experimental	32	0.23	32	0.12
Narrative Intelligence Scale posttest	Control	32	0.23	32	0.73
	Experimental	32	0.13	32	0.16

3. Results

The results of the independent sample *t*-test performed for the narrative intelligence research problem is shown in Table 6. When Table 6 is examined, it can be seen that there is a significant difference in favor of the experimental group in terms of undergraduate students' narrative intelligence scores ($p < 0.05$). Therefore, the result is that the narrative intelligence of the students in the experimental group is higher. Additionally, Cohen's d was used to give an idea of how significant the differences between the groups are. According to the results in Table 6, the sizes of the differences between groups for each measure vary from medium to large.

Table 6. *T*-test results of narrative intelligence scores of each group.

	Group	N	M	SD	df	t	*p*	Cohen's d
Emplotment	Control	32	4.52	0.72	62	−2.57	0.02 *	0.61
	Experimental	32	4.91	0.66				
Characterization	Control	32	4.55	0.71	62	−2.87	0.01 *	0.74
	Experimental	32	4.88	0.62				
Generation	Control	32	4.22	0.63	62	−2.74	0.01 *	0.71
	Experimental	32	4.53	0.66				
Narration	Control	32	4.66	0.65	62	−2.64	0.01 *	0.73
	Experimental	32	4.84	0.64				
Thematization	Control	32	4.44	0.77	62	−3.61	0.00 *	0.93
	Experimental	32	4.88	0.61				

* $p < 0.001$.

Table 6 presents the results of *t*-tests comparing the control and experimental group scores in various dimensions of narrative intelligence. The first dimension, "Emplotment," shows that the experimental group significantly outperformed the control group, with a mean score of 4.91 compared to 4.52. The *t*-test results reveal a significant difference (t = −2.57, *p* = 0.02), suggesting that the intervention applied to the experimental group had a notable impact on their "Emplotment" ability, with a moderate effect size (Cohen's d = 0.61). With regards to the "Characterization" dimension, the control group once again displayed higher performance, with a mean score of 4.55 compared to the experimental group's 4.88. The *t*-test indicates a significant difference (t = −2.87, *p* = 0.01) and a moderate-to-large effect size (Cohen's d = 0.74), emphasizing the effectiveness of the intervention in improving characterization skills. Effect size measures, such as Cohen's d, were employed to gauge the extent of the differences between the control and experimental groups. This approach aids readers in comprehending the real-world significance of the observed enhancements [33,34]. Similarly, in "Generation" and "Narration," the experimental group consistently outperformed the control group, as supported by significant *t*-test results, with *p* = 0.01. Effect sizes (Cohen's d) for these dimensions were moderate, with

values of 0.71 and 0.73, respectively. These findings suggest that the experimental group exhibited a clear advantage in generating narratives and storytelling. The most striking contrast emerged in the "Thematization" dimension, where the experimental group achieved a significantly higher mean score (M = 4.44) compared to the control group (M = 4.88), with a highly significant *t*-test result (t = −3.61, *p* = 0.00) and a substantial effect size (Cohen's d = 0.93). This underscores the substantial impact of the intervention on thematic development skills. In summary, the *t*-test results consistently indicate that the experimental group exhibited significant improvements in various facets of narrative intelligence compared to the control group. The effect sizes range from moderate to large, underscoring the effectiveness of the applied intervention across these dimensions. These findings provide valuable insights into the positive impact of the treatment on narrative intelligence.

Table 7 presents the undergraduates' mean SAWSES scores in both the pretests and posttests. The table presents mean and standard deviation values for the different categories of the SAWSES across two time points: "pretest" and "posttest". The two groups, the "control" group and the "experimental" group, are compared within each category. Notably, the experimental group consistently exhibited higher mean SAWSES scores in the posttest compared to the control group, indicating potential improvements in self-assessment and self-efficacy. Across all categories, mean scores increased from the pretest to the posttest for both groups, suggesting that the intervention may have positively impacted their self-assessment in these areas. To draw more robust conclusions and ascertain the statistical significance of these changes, further statistical analysis would be necessary, considering the research context and specific hypotheses.

Table 7. Mean and standard deviation values of the SAWSES scale.

SAWSES	Group	Pretest			Posttest		
		N	**M**	**SD**	**N**	**M**	**SD**
Writing essentials	Control	32	3.33	0.55	32	3.67	0.66
	Experimental	32	3.53	0.56	32	4.13	0.56
Reflective writing	Control	32	3.17	0.62	32	3.57	0.67
	Experimental	32	3.42	0.45	32	3.88	0.52
Creative identity	Control	32	3.64	0.78	32	3.78	0.79
	Experimental	32	3.88	0.67	32	4.38	0.62
General writing self-efficacy	Control	32	3.33	0.49	32	3.78	0.68
	Experimental	32	3.64	0.59	32	3.93	0.64

Table 7 provides an insightful look into the evolution of the students' self-assessment and self-efficacy in writing. In the "Writing essentials" category, both groups displayed improvements in their self-assessment. Notably, the experimental group exhibited a more substantial increase in their mean SAWSES score, from 3.53 in the pretest to 4.13 in the posttest, suggesting that the intervention had a noteworthy positive impact on their self-assessment. A similar trend is observed across other categories, such as "Reflective writing" and "General writing self-efficacy," where the experimental group consistently outperformed the control group, indicating that the intervention likely contributed to their improved self-assessment and self-efficacy. The "Creative identity" indicator stands out as a category where the control group's mean SAWSES score remained relatively stable, while the experimental group demonstrated significant improvement. This further underscores the effectiveness of the intervention in nurturing a creative identity in the participants. In summary, the findings imply that the experimental group experienced more significant enhancements in self-assessment and self-efficacy over time, across various aspects of writing, compared to the control group. These results suggest that the intervention or treatment implemented for the experimental group had a beneficial effect on their self-assessment and self-efficacy in writing-related areas. Further statistical analyses would be advisable to determine the statistical significance of these observed improvements.

Calculated effect size measures (e.g., Cohen's d) were additionally utilized to quantify the magnitude of the differences between the control and experimental groups. This can help readers understand the practical significance of the observed improvements [33,34]. In the "Writing essentials" dimension, the effect size (Cohen's d) is approximately 0.75, which signifies a moderately large difference between the experimental and control groups. This suggests that the experimental group experienced a substantial improvement in their self-assessment and self-efficacy related to writing essentials compared to the control group. The effect size of 0.75 reflects a notable impact of the intervention in this dimension, highlighting the significance of the treatment in enhancing self-assessment and self-efficacy in this specific aspect of writing. In the "Reflective writing" dimension, the effect size is about 0.52, indicating a moderate difference between the experimental and control groups. This implies that after the intervention, the experimental group demonstrated a moderate improvement in self-assessment and self-efficacy in terms of reflective writing compared to the control group. While the effect size is not as large as in the "Writing essentials" dimension, it still underscores the positive impact of the intervention on this specific dimension of writing. The "Creative identity" dimension exhibits a substantial effect size of approximately 0.85, signifying a large difference between the experimental and control groups. This result suggests that the intervention had a profound impact on nurturing a creative identity in the experimental group. Participants in the experimental group showed significantly higher self-assessment and self-efficacy in this dimension compared to the control group. The substantial effect size emphasizes the effectiveness of the treatment in fostering a creative identity among the participants. In the "General writing self-efficacy" dimension, the effect size is about 0.23, which indicates a small-to-moderate difference between the experimental and control groups.

While the experimental group did exhibit an improvement in self-assessment and self-efficacy for writing self-efficacy in general, the effect size is not as pronounced as in the other dimensions. It suggests that the intervention had a smaller, yet noticeable, impact on this specific aspect of writing self-efficacy. For this reason, the results of this study demonstrate that the intervention had varying effects across the four dimensions of self-assessment and self-efficacy in writing. The "Creative Identity" dimension showed the most significant improvement, followed by "Writing essentials" and "Reflective writing", both of which demonstrated moderate improvements. However, the "General writing self-efficacy" dimension showed a smaller, albeit discernible, improvement. These effect sizes provide valuable insights into the practical significance of the intervention in terms of enhancing self-assessment and self-efficacy in different facets of writing.

When Table 8 is examined, it can be seen that there was a significant difference in favor of the experimental group in terms of SAWSES scores ($F(1, 62) = 4.548$, $p = 0.039$), the "Writing essentials" factor ($F(1, 62) = 4.792$, $p = 0.037$), and the "Reflective writing" factor ($F(1, 62) = 6.846$, $p = 0.016$) ($p < 0.05$). However, there was no significant difference in the "Creative identity" factor ($F(1, 62) = 0.007$, $p = 0.968$) ($p > 0.05$). There are differences within the groups and between the groups in terms of writing self-efficacy ($p < 0.05$). These differences are in favor of the experimental group. In general, the intervention was found to increase the writing self-efficacy of undergraduate students. According to the results in Table 8, the sizes of inter-group difference for each measure vary from medium to large (η^2partial = 0.08).

Table 8 provides a comprehensive overview of the outcomes of the analysis of variance (ANOVA) conducted on the pretest and posttest SAWSES scores. In the "Writing essentials" category, the ANOVA reveals a significant difference between the control and experimental groups in terms of their pretest and posttest scores, with a notable $F = 6.129$ and a $p = 0.018$. This suggests that the intervention or treatment had a meaningful impact on participants' self-assessment related to writing essentials, with a moderate effect size (partial eta squared $\eta^2 = 0.09$). Similarly, "Reflective writing" demonstrates a substantial difference between the groups, as indicated by a high $F = 8.693$ and a $p = 0.005$. This emphasizes the impact of

the intervention in enhancing participants' reflective writing skills, with a moderate effect size ($\eta^2 = 0.12$).

Table 8. ANOVA results for the SAWSES scale (pretest and posttest scores).

Variable	Source	Sum of Squares	df	Mean Square	F	p	Partial Eta Squared (η^2)
Writing essentials	Group	3.089	1	3.197	6.129 *	0.018	0.09
	Error	32.998	62	0.524			
	Scores	0.762	1	0.768	8.698 *	0.004	0.13
	Group * Score	0.388	1	0.389	4.792 *	0.037 *	0.08
	Error	5.287	62	0.091			
	Total	41.552	127				
Reflective writing	Group	4.698	1	4.762	8.693 *	0.005	0.12
	Error	35.875	62	0.555			
	Scores	1.974	1	1.778	23.512 *	0.000	0.27
	Group * Score	0.566	1	0.569	6.846 *	0.016 *	0.10
	Error	4.936	62	0.079			
	Total	46.867	117				
Creative identity	Group	2.298	1	2.368	2.790	0.113	0.04
	Error	55.348	62	0.866			
	Scores	0.936	1	0.911	3.995	0.056	0.06
	Group * Score	0.003	1	0.003	0.007	0.968	0.00
	Error	14.698	62	0.241			
	Total	73.241	117				
General writing self-efficacy	Group	4.293	1	3.197	6.54 *	0.017	0.09
	Error	4.624	62	0.521			
	Scores	0.991	1	0.984	14.559 *	0.000	0.19
	Group * Score	0.308	1	0.298	4.548 *	0.039 *	0.07
	Error	4.298	62	0.078			
	Total	40.469	117				

* $p < 0.001$.

However, "Creative identity" did not display a significant difference between groups, with a *p*-value of 0.113 and a negligible effect size ($\eta^2 = 0.04$). This suggests that the treatment did not significantly influence creative identity in this context. In the "General writing self-efficacy" category, the ANOVA results point to a significant impact of the intervention, with an F = 6.54 and a *p* = 0.017. This underscores the efficacy of the intervention in improving participants' general writing self-efficacy, with a moderate effect size ($\eta^2 = 0.09$). Overall, the ANOVA results highlight the differential impact of the intervention on various aspects of self-assessment and self-efficacy related to writing. While it significantly improved "Writing essentials", "Reflective writing", and "General writing self-efficacy", it had no substantial effect on "Creative identity". These findings provide valuable insights into the effectiveness of the treatment in different facets of writing self-assessment and self-efficacy.

4. Discussion and Conclusions

The current study delves into the effects of engaging in digital storytelling using different web-based platforms by undergraduate students majoring in Primary Education with a twofold purpose: (a) it investigates the extent to which participating in digital story creation activities influences the narrative intelligence of undergraduate students and (b) it explores whether narrative intelligence in these digital tasks impacts their writing self-efficacy. The results show that the experimental group, which used generative AI platforms, had significantly higher scores in both measures than the control group, which

used traditional platforms. These findings suggest that generative AI platforms have the potential to improve undergraduates' narrative and writing self-efficacy in fostering story creation [22,24].

Regarding *RQ1*, the results aimed to investigate whether the proposed treatment had a significant impact on the narrative intelligence scores of undergraduate students. The findings reveal a noteworthy difference in favor of the experimental group. In line with previous studies [8,21], this study's findings indicate that the levels of narrative intelligence among undergraduate students exposed to the experimental AI-supported instructional method were significantly higher than those of the control group. Furthermore, the effect sizes, as indicated by Cohen's d, ranged from medium to large. These effect sizes emphasize the practical significance of the observed differences, further supporting the claim that the experimental method positively influenced students' narrative intelligence [12,14]. Furthermore, *t*-test results comparing narrative intelligence dimensions in the control and experimental groups showed that for the "Emplotment" indicator, the experimental group significantly outperformed the control group (t = −2.57, p = 0.02, Cohen's d = 0.61). "Characterization" also favored the participants from the former group (t = −2.87, p = 0.01, Cohen's d = 0.74). Similarly, "Generation" and "Narration" showed significant improvements (p = 0.01) with moderate effect sizes (Cohen's d = 0.71 and Cohen's d = 0.73 to each, respectively). The most substantial contrast was in "Thematization" (t = −3.61, p = 0.00, Cohen's d = 0.93). Hence, *t*-tests consistently highlighted significant improvements in narrative intelligence, with moderate-to-large effect sizes, affirming the intervention's positive impact.

In relation to *RQ2*, the analysis of writing self-efficacy, measured through the SAWSES, was conducted using a repeated measurements two-way ANOVA. Consistent with previous studies' findings [4,19], the evidence from this comparative analysis allowed us to explore whether the proposed treatment had a significant impact on the writing self-efficacy of undergraduate students. The analysis also indicated a significant difference in favor of the experimental group in the "Writing essentials" factor. This suggests that the proposed treatment was effective in improving students' understanding of fundamental writing principles, further supporting the positive impact of the experimental approach. In the "Reflective writing" factor, a significant difference was observed in favor of the experimental group. In the "Reflective writing" category, ANOVA results similarly indicated a substantial impact of the intervention. Effective reflective writing is essential in academic and professional settings, and this result underscores the potential for interventions to positively influence such skills. This result suggests that the proposed treatment encouraged students to engage in reflective and relational writing practices, contributing to their overall writing self-efficacy [11,16]. Nonetheless, it is worth noting that there was no significant difference in the "Creative identity" factor. This implies that the proposed treatment did not have a significant impact on students' perceptions of their creative identity in writing. This study's outcomes highlight the importance of tailored interventions in the development of writing skills and their self-assessment. Educators and researchers should consider the distinct nature of these dimensions and develop strategies that align with their specific goals and objectives.

Overall, the findings demonstrate a notable distinction in favor of the experimental group concerning overall writing self-efficacy scores. This observation indicates that the use of generative AI platforms in literacy education had a favorable impact on bolstering undergraduate students' writing self-efficacy. The medium-to-large effect size underscores the practical importance of this outcome. This comparative study sheds light on the effectiveness of digital storytelling in various instructional design contexts and provides insights into the importance of the variances observed between the experimental and control groups. The results of this research carry significance not only for the academic community but also for educational practitioners and technology developers. Understanding whether digital story creation activities influence narrative intelligence and writing self-efficacy can offer valuable insights into the design and implementation of generative AI platforms to enhance, rather than impede, individuals' storytelling skills and confidence in the digital

age [17]. This study contributes to a more comprehensive understanding of the evolving convergence of education, technology, and narrative intelligence, providing a foundation for further exploration in the field of literacy education.

In conclusion, this study's findings indicate that the use of generative AI platforms in the context of story creation tasks can substantially enhance both narrative intelligence scores and writing self-efficacy when compared to conventional. These results also underscore the potential of generative AI platforms in educational settings, especially in elevating students' competencies and confidence in narrative composition. While the study demonstrated notable improvements in narrative intelligence and writing self-efficacy, it is worth mentioning that no statistically distinguishable distinctions were observed in the "Creative identity" aspect. This suggests that AI technologies may excel in specific facets of writing development but may not entirely replace or replicate the creative dimensions of human storytelling.

5. Implications

The present study contributes to the corpus of knowledge of how technology can be harnessed to enhance writing skills and confidence in future educators. It highlights the potential of generative AI platforms as valuable tools in educational settings, opening up opportunities for further exploration and integration of technology to support students in their writing endeavors. These findings have important implications for educators, curriculum designers, and institutions seeking to optimize the use of web-based AI platforms in fostering narrative intelligence and writing self-efficacy among undergraduate students.

This study's findings have several implications for educational practice. Generative AI platforms can foster narrative intelligence and writing self-efficacy in undergraduate Primary Education majors, even among those with limited computer science skills. Second, they suggest that educators should consider integrating generative AI platforms into their teaching practices. Third, the results of this study suggest that future research should investigate the long-term effects of using generative AI platforms on undergraduates' narrative and writing skills. Therefore, the implications of this study's findings, considering their relevance in the context of contemporary education and the broader field of digital storytelling and generative AI integration, are as follows:

1. Enhancement of narrative intelligence and writing self-efficacy: This study's results provide evidence that the use of generative AI platforms can significantly improve undergraduate students' narrative intelligence scores and writing self-efficacy. This finding aligns with the growing recognition of the potential benefits of technology-assisted learning in enhancing core competencies. Educators and institutions can leverage these tools to empower students with the skills and confidence necessary for effective communication;

2. Effects of generative AI platforms on storytelling: The positive outcomes observed in the experimental group highlight the evolving role of AI technologies in education. Generative AI platforms can augment traditional teaching methods by offering personalized feedback, suggesting improvements, and facilitating the writing process. The study suggests that these technologies can serve as valuable educational aids, supporting students in their journey to become proficient writers;

3. Human creativity vs. generative AI assistance: Educators and curriculum designers should consider incorporating generative AI platforms into writing instruction. On the one hand, while generative AI demonstrated clear benefits in terms of narrative intelligence and writing self-efficacy, the absence of a significant difference in the "Creative identity" factor suggests that human creativity remains a distinctive and irreplaceable aspect of storytelling. This finding underscores the importance of striking a balance between generative AI assistance and human creativity, particularly in fields where originality and creative expression are highly valued. On the other hand, it is crucial to do so thoughtfully, recognizing the strengths and limitations of AI tools. These platforms

can be especially useful for tasks involving grammar, structure, and organization, leaving space for students to focus on the creative aspects of their writing.

6. Limitations and Future Work

The present study encountered certain limitations worth acknowledging. First, this study's sample size was relatively small, with only 64 undergraduate students from a Greek department participating. This limited sample size may not be representative of the entire undergraduate population, and caution should be exercised when generalizing the results to a larger and more diverse population of students. This study's sample also comprised only undergraduate students from a single primary education department. As a result, there were notable variations in the socio-cognitive backgrounds of the participants, which may limit the generalizability of the findings to a more diverse population. Second, while the study compared the effects of different digital platforms, it was primarily focused on well-known platforms.

From a methodological research perspective, the absence of interview data to complement the quantitative findings is a noteworthy limitation. While an analysis of learning behaviors helped to provide context for the quantitative results, the lack of qualitative data obtained through interviews may have limited the depth of their understanding. Third, this study had a short-term perspective, as it examined the immediate impact of digital story creation tasks on narrative intelligence and writing self-efficacy. Long-term effects and whether the observed improvements are sustained over time were not explored. Fourth, this study primarily focused on undergraduate students, and the results may not apply to other educational levels or professional contexts. Fifth, self-report methods were used to collect data on participants' thoughts, feelings, and behaviors. However, it is important to note that self-reporting methods are subject to a variety of biases, such as social desirability bias, memory bias, introspection bias, and demand characteristics. These biases could have affected the accuracy of the results. It would be helpful to use multiple methods of data collection, such as interviews and observations, to triangulate the findings and reduce the impact of biases.

These findings open avenues for further research in the integration of AI in education. Future studies could explore the long-term effects of using generative AI platforms on students' writing skills and assess the transferability of these skills to real-world writing contexts. The above limitations suggest that the following future works are worth pursuing:

1. Long-term effects: Conducting longitudinal studies to assess the long-term impact of digital storytelling and generative AI technologies on narrative intelligence and writing self-efficacy. This would provide a more comprehensive understanding of the sustained benefits or potential drawbacks of using these tools;
2. Diversity and inclusion: Investigating how digital storytelling and AI technologies affect students from diverse backgrounds, including those with varying levels of writing proficiency, linguistic diversity, and accessibility needs, would ensure that the benefits are accessible to a broad range of learners;
3. Pedagogical strategies: Exploring different pedagogical approaches and instructional designs that maximize the benefits of digital storytelling and AI technologies in educational contexts. Developing guidelines and best practices for educators to effectively integrate these tools into their teaching;
4. Interdisciplinary research: Collaborating with experts from the fields of psychology, education, and computer science to gain a more holistic understanding of the cognitive processes involved in digital storytelling and generative AI writing. This interdisciplinary approach can shed light on the underlying mechanisms at play;
5. Ethical considerations: Investigating the ethical implications of using AI in education, especially regarding issues such as plagiarism detection, privacy, and potential bias in AI-generated content. Developing ethical guidelines for the responsible use of AI technologies in educational settings.

References

1. Rodríguez, C.C.; Romero, C.N. *Digital Storytelling in Primary Education: A Comparison of Already Made Resources and Story Creation Tools to Improve English as a Foreign Language*; Octaedro: Barcelona, Spain, 2023; p. 387.
2. Nurlaela Ilham, M.J.; Lisabe, C. The Effects of Storyjumper on Narrative Writing Ability of EFL Learners in Higher Education. *J. Educ. FKIP UNMA* **2022**, *8*, 1641–1647.
3. Güvey, A. Writing a Folktale as an Activity of Written Expression: Digital Folktales with Storyjumper. *Educ. Policy Anal. Strateg. Res.* **2020**, *15*, 159–185. [CrossRef]
4. Dai, J.; Wang, L.; He, Y. Exploring the Effect of Wiki-based Writing Instruction on Writing Skills and Writing Self-efficacy of Chinese English-as-a-foreign Language Learners. *Front. Psychol.* **2023**, *13*, 1069832. [CrossRef]
5. Ispir, B.; Yıldız, A. An Overview of Digital Storytelling Studies in Classroom Education in Turkey. *J. Qual. Res. Educ.* **2023**, *23*, 187–216. [CrossRef]
6. Zakaria, S.M.; Yunus, M.M.; Nazri, N.M.; Shah, P.M. Students' Experience of Using Storybird in Writing ESL Narrative Text. *Creative Educ.* **2016**, *7*, 2107–2120. [CrossRef]
7. Fitriyani, R.; Ahsanu, M.; Kariadi, M.T.; Riyadi, S. Teaching Writing Skills through Descriptive Text by Using Digital Storytelling "storyjumper". *J. Engl. Educ. List.* **2023**, 4.
8. Fang, X.; Ng, D.T.K.; Leung, J.K.L.; Chu, S.K.W. A Systematic Review of Artificial Intelligence Technologies Used for Story Writing. *Educ. Inf. Technol.* **2023**, 1–37. [CrossRef]
9. Lambert, J.; Stevens, M. Chatgpt and Generative AI Technology: A Mixed Bag of Concerns and New Opportunities. *Comput. Sch.* **2023**, 1–25. [CrossRef]
10. Han, A.; Cai, Z. Design Implications of Generative AI Systems for Visual Storytelling for Young Learners. In Proceedings of the 22nd Annual ACM Interaction Design and Children Conference, Chicago, IL, USA, 19–23 June 2023; pp. 470–474.
11. O'Meara, J.; Murphy, C. Aberrant AI Creations: Co-creating Surrealist Body Horror Using the DALL-E Mini Text-to-image Generator. *Converg. Int. J. Res. New Media Technol.* **2023**, *29*, 1070–1096. [CrossRef]
12. Bender, S.M. Coexistence and Creativity: Screen Media Education in the Age of Artificial Intelligence Content Generators. *Media Pract. Educ.* **2023**, 1–16. [CrossRef]
13. Chaudhary, S. 15 Best Free AI Content Generator & AI Writers for 2023. 2023. Available online: https://medium.com/@bedigisure/free-ai-content-generator-11ef7cbb2aa0 (accessed on 10 October 2023).
14. Lee, J.H.; Shin, D.; Noh, W. Artificial Intelligence-based Content Generator Technology for Young English-as-a-foreign-language Learners' Reading Enjoyment. *RELC J.* **2023**, *54*, 508–516. [CrossRef]
15. Im, H.; Jeon, S.; Cho, H.; Shin, S.; Choi, D.; Hong, H. AI Story Relay: A Collaborative Writing of Design Fiction to Investigate Artificial Intelligence Design Considerations. In Proceedings of the 2023 ACM International Conference on Supporting Group Work, Hilton Head, SC, USA, 8–11 January 2023; pp. 6–8.
16. Pellas, N. The influence of sociodemographic factors on students' attitudes toward AI-generated video content creation. *Smart Learn. Environ.* **2023**, *10*, 57. [CrossRef]
17. Sun, T.; Wang, C. College Students' Writing Self-efficacy and Writing Self-regulated Learning Strategies in Learning English as a Foreign Language. *System* **2020**, *90*, 102221. [CrossRef]
18. Teng, M.F.; Wang, C. Assessing Academic Writing Self-Efficacy Belief and Writing Performance in a Foreign Language Context. *Foreign Lang. Ann.* **2023**, *56*, 144–169. [CrossRef]
19. Campbell, C.W.; Batista, B. To Peer or Not to Peer: A Controlled Peer-editing Intervention Measuring Writing Self-efficacy in South Korean Higher Education. *Int. J. Educ. Res. Open* **2023**, *4*, 100218. [CrossRef]
20. Randall, W.L. Narrative intelligence and the novelty of our lives. *J. Aging Stud.* **1999**, *13*, 11–28. [CrossRef]
21. Ziaei, S.; Ghonsooly, B.; Ghabanchi, Z.; Shahriari, H. Proposing a Cognitive EFL Writing Model Based on Personality Types and Narrative Writing Intelligence: A SEM Approach. *Issues Lang. Teach.* **2019**, *8*, 163–185.
22. Golabi, A.; Heidari, F.F. Prediction of Iranian EFL Learners' Learning Approaches Through Their Teachers' Narrative Intelligence and Teaching Styles: A Structural Equation Modelling Analysis. *Iran. J. Engl. Acad. Purp.* **2019**, *8*, 1–12.
23. Pishghadam, R.; Baghaei, P.; Shams, M.A.; Shamsaee, S. Construction and Validation of a Narrative Intelligence Scale with the Rasch Rating Scale Model. *Int. J. Educ. Psychol. Assess.* **2011**, *8*, 75–90.
24. Ge, J.; Lai, J.C. Artificial Intelligence-based Text Generators in Hepatology: ChatGPT Is Just the Beginning. *Hepatol. Commun.* **2023**, *7*, e0097. [CrossRef]

25. Yuan, A.; Coenen, A.; Reif, E.; Ippolito, D. Wordcraft: Story Writing with Large Language Models. In Proceedings of the 27th International Conference on Intelligent User Interfaces, Helsinki, Finland, 22–25 March 2022; pp. 841–852.

26. Pishghadam, R.; Shams, M.A. Hybrid Modeling of Intelligence and Linguistic Factors as Predictors of L2 Writing Quality: A SEM Approach. *Lang. Test. Asia* **2012**, *2*, 1–24. [CrossRef]

27. Cohen, L.; Manion, L.; Morrison, K. *Research Methods in Education*; Routledge: New York, NY, USA, 2007.

28. Rogers, J.; Revesz, A. *Experimental and Quasi-Experimental Designs*; Routledge: London, UK, 2019; pp. 133–143.

29. Cennamo, K.; Ross, J.; Ertmer, P. *Technology Integration for Meaningful Classroom Use. A Standard-based Approach*; Wadsworth: Belmont, CA, USA, 2010.

30. Collins, A.; Joseph, D.; Bielaczyc, K. *Design Research: Theoretical and Methodological Issues*; Psychology Press: Portland USA, 2016.

31. Mitchell, K.M.; McMillan, D.E.; Lobchuk, M.M.; Nickel, N.C.; Rabbani, R.; Li, J. Development and Validation of the Situated Academic Writing Self-Efficacy Scale (SAWSES). *Assess. Writ.* **2021**, *48*, 100524. [CrossRef]

32. Cortina, J.M. What Is Coefficient Alpha? an Examination of Theory and Applications. *J. Appl. Psychol.* **1993**, *78*, 98–104. [CrossRef]

33. Garth, A. *Analyzing Data Using SPSS*; Sheffield Hallam University: Sheffield, UK, 2008.

34. Cha, E.-S.; Kim, K.H.; Erlen, J.A. Translation of Scales in Cross-cultural Research: Issues and Techniques. *J. Adv. Nurs.* **2007**, *58*, 386–395. [CrossRef] [PubMed]