Is ChatGPT helpful for graduate students in acquiring knowledge about digital storytelling and reducing their cognitive load? An experiment.

¿ChatGPT es útil para que los estudiantes de posgrado adquieran conocimientos sobre narración digital y reduzcan su carga cognitiva? Un experimento.

Raidell Avello-Martínez
Universidad Bolivariana del Ecuador, Guayaquil, Ecuador
ravellom@ube.edu.ec

Tomasz Gajderowicz
Universidad de Varsovia, Varsovia, Polonia
tgajderowicz@wne.uw.edu.pl

Víctor G. Gómez-Rodríguez
Universidad Bolivariana del Ecuador, Guayaquil, Ecuador
vgomez@ube.edu.ec

Abstract

This study examines the impact of ChatGPT on narrative scriptwriting abilities and cognitive load in a sample of 41 master's students enrolled in a Digital Narratives course. Using a randomized experimental design, participants were divided into two groups: an experimental group (n = 20) that interacted with ChatGPT and a control group (n = 21) that did not. Our methods involved pre- and post-tests to assess changes in digital storytelling skills and cognitive load, as defined by intrinsic, extraneous, and germane load measures. The results indicated no significant improvement in digital storytelling skills for the experimental group compared to the control group, suggesting that the use of ChatGPT does not markedly enhance narrative writing abilities in the short term. However, a significant reduction in germane cognitive load was observed among the experimental group, pointing to ChatGPT's potential to facilitate the learning process by reducing the mental effort required for task integration and application. The study underscores the complexity of integrating AI into learning environments and highlights the need for strategic AI implementation tailored to specific educational objectives. It also points to the importance of longitudinal research to fully understand the long-term effects of AI on learning and cognitive development.

Keywords: Artificial Intelligence in Education, Digital Storytelling, Cognitive Load Theory, ChatGPT, Narrative Writing Skills.

Resumen

Este estudio examina el impacto de ChatGPT en las habilidades de escritura de guiones narrativos y la carga cognitiva en una muestra de 41 estudiantes de maestría matriculados en un curso de Narrativas Digitales. Utilizando un diseño experimental aleatorio, los participantes se dividieron en dos grupos: un grupo experimental (n = 20) que interactuaba con ChatGPT y un grupo de control (n = 21) que no lo hacía. Nuestros métodos incluyeron pruebas previas y posteriores para evaluar los cambios en las habilidades de narración digital y la carga

cognitiva, según lo definido por medidas de carga intrínsecas, extrañas y pertinentes. Los resultados no indicaron una mejora significativa en las habilidades de narración digital para el grupo experimental en comparación con el grupo de control, lo que sugiere que el uso de ChatGPT no mejora notablemente las habilidades de escritura narrativa a corto plazo. Sin embargo, se observó una reducción significativa en la carga cognitiva relevante entre el grupo experimental, lo que apunta al potencial de ChatGPT para facilitar el proceso de aprendizaje al reducir el esfuerzo mental requerido para la integración y aplicación de tareas. El estudio subraya la complejidad de integrar la IA en entornos de aprendizaje y destaca la necesidad de una implementación estratégica de la IA adaptada a objetivos educativos específicos. También señala la importancia de la investigación longitudinal para comprender plenamente los efectos a largo plazo de la IA en el aprendizaje y el desarrollo cognitivo.

Palabras clave: Inteligencia Artificial en Educación, Narración Digital, Teoría de la Carga Cognitiva, ChatGPT, Habilidades de Escritura Narrativa.

1. Introduction

In the digital age, narrative has undergone a profound transformation, evolving from its traditional roots towards forms that leverage digital tools and multimedia formats to enrich the storytelling experience. This paradigm shift has given rise to what we now recognize as digital narratives, a contemporary form of art that allows creators to seamlessly integrate multimedia elements, crafting tales that are both captivating and immersive (Hurtado-Mazeyra et al., 2023; Rodriguez et al., 2021). Digital narratives, due to their versatile nature, come in various forms and are delivered through a wide range of digital platforms, employing words, images, videos, audio, and interactive elements (Wu & Chen, 2020). This evolution has not only expanded the boundaries of creativity but also facilitated the development of writing skills, creativity, and digital literacy in students, preparing them to effectively participate in the contemporary digital landscape (Avello-Martínez et al., 2023).

Parallelly, the integration of artificial intelligence (AI) in the educational sphere has sparked significant academic and pedagogical debate (Essel et al., 2024; García-Peñalvo, 2023). Among emerging technologies, AI-based language models, like ChatGPT, have gained prominence due to their ability to generate responses that closely mimic human interactions. These tools have revolutionized the way we interact with machines, offering new opportunities for educational support and content generation (Bai et al., 2023). In this context, ChatGPT's ability to facilitate writing and narrative suggests considerable potential for its application in higher education, particularly in disciplines that require a high degree of innovation, critical thinking, and creativity (Pellas, 2023).

Digital storytelling, merging the ancient art of storytelling with modern technology, offers a fertile ground for educational innovation. The adoption of generative AI platforms in education proposes a fundamental shift in how educators and students approach narrative creation (Pellas, 2023). These tools allow for exploring new ways of storytelling, transcending the limits of written text to include multimedia resources that enrich the narrative (Fang et al., 2023). This enhancement of narrative practices not only improves students' writing and

creative expression skills but also fosters a deeper and more critical understanding of digital technologies and their potential for effective communication.

The inclusion of ChatGPT in higher education has been analyzed for its potential to transform the educational experience, enhancing interaction and learning through a conversational platform that promotes dynamic idea exchange. ChatGPT's ability to generate coherent and contextually relevant dialogues, as well as to support the creation of complex characters and plots, indicates its value as an educational tool in teaching narrative writing (Essel et al., 2024). Beyond facilitating content generation, ChatGPT's interactive nature can promote reflective and deep learning, enabling students to question, clarify, and connect information with their prior knowledge, thereby fostering more meaningful learning. Additionally, it can also help learners accurately manage their own cognitive load (Vandewaetere & Clarebout, 2013).

However, despite initial enthusiasm, the implementation of generative AI technologies in education is still in its early stages, with numerous questions about their impact on the learning experience. Assessing these tools in terms of effectiveness, safety, and ethics is crucial for their responsible and effective integration into educational processes. Moreover, understanding students' perceptions of these technologies and how they influence their learning process and collaboration with the machine is essential. This is particularly relevant at a time when higher education faces unprecedented changes and opportunities derived from the integration of technology into the classroom.

Within the broader discourse on the integration of digital narratives and artificial intelligence tools like ChatGPT in education, cognitive load theory (CLT) becomes an indispensable theoretical foundation for evaluating the effectiveness of such technological advances (van Gog et al., 2010). CLT, a cornerstone in the educational psychology landscape, underscores the importance of tailoring learning experiences to align with human cognitive architecture. It argues that for learning materials to be truly effective, they must be designed with a deep understanding of the cognitive load they impose on students. This involves a nuanced differentiation between intrinsic, extraneous, and relevant cognitive loads, each of which plays a distinct role in the learning process (Klepsch et al., 2017).. The arrival of instruments capable of measuring these specific cognitive loads marks a significant advance in educational research, offering a refined lens through which the impact of educational interventions, including those involving intelligence technologies, can be more accurately assessed. artificial like ChatGPT. These tools promise not only to improve instructional design, but also to elevate the overall quality of educational experiences by ensuring that learning is aligned with cognitive principles (Klepsch et al., 2017).

The present study focuses on exploring the impact of ChatGPT use on narrative script writing and the cognitive load of master's students in educational technology. Through an experimental approach, it seeks to quantitatively assess how interaction with ChatGPT can serve as support in developing specific narrative writing skills for educational videos and whether this tool can contribute to reducing the cognitive load associated with this creative process. Cognitive load refers to the amount of mental resources required to perform a task; in the context of creative writing, a tool that can decrease this load could facilitate a more efficient and enjoyable learning process for students.

This study aims to provide a deeper understanding of how AI tools, specifically ChatGPT, can influence the learning process and the quality of written productions in educational

contexts. By exploring this issue, it aspires to contribute to the body of knowledge on the application of AI tools in education, addressing both the possibilities these technologies offer to enrich learning and teaching and the challenges their integration presents. Research on the interaction between students and AI technologies is fundamental to understanding not only the potential of these tools to enhance the educational process but also to identify and address the potential risks and limitations that their use may entail.

Hypothesis development

The literature offers varied views on ChatGPT's role in enhancing writing, pointing to the need for a quantitative approach. Accordingly, this study proposes hypotheses to explore ChatGPT's effect on digital storytelling, aiming for definitive insights:

- **H1** (Baseline Hypothesis): There is no significant difference in digital storytelling knowledge between the experimental and control groups at baseline, confirming initial group equivalence.
- **H2**: Participants will show a significant improvement in digital storytelling scores from pre-test to post-test within each group, indicating ChatGPT's educational utility.
- **H3**: There will be a significant difference in digital storytelling abilities, specifically in character development, plot structuring, and dialogue crafting, between the experimental and control groups, underscoring ChatGPT's influence on creative writing components.
- **H4**: Cognitive load differences between groups post-intervention will reveal ChatGPT's effectiveness in reducing the mental effort required in the storytelling process.

2. Methodology

Study Design

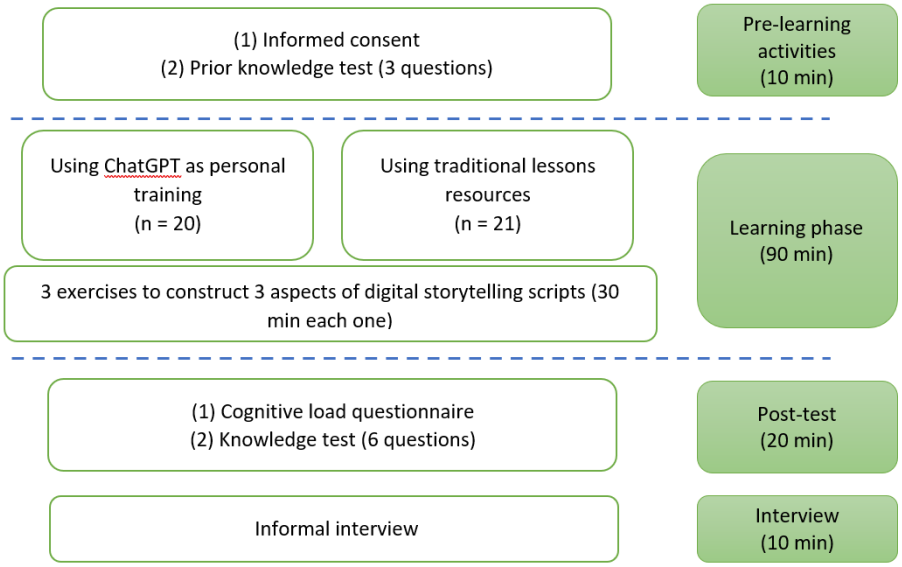
This study adopted an experimental design (pre-test, post-test) to investigate the impact of using ChatGPT on the acquisition of knowledge about digital storytelling and the reduction of cognitive load in master's students. A total of 41 students enrolled in the online Digital Narratives course were divided into two groups: an experimental group and a control group, using Google Meet's random group creation feature to ensure random assignment.

Participants

The study involved 41 master's students enrolled in the Digital Narratives course. Participants were randomly assigned to one of two groups: the experimental group using ChatGPT (n = 20) and the control group (n = 21).

Procedure

The study was conducted in three main phases: pre-learning activities, learning phase, and post-intervention assessment.



Pre-learning activities (10 minutes): All students completed the following activities:

- Signing of informed consent.
- A prior knowledge test consisting of 3 questions to assess the initial level of knowledge about digital storytelling, using Moodle questionnaire.

Learning phase (90 minutes):

- **Group 1 (Experimental):** This group used ChatGPT as a personalized learning support tool. They were assigned three exercises aimed at constructing three essential aspects of digital storytelling scripts: character creation, plot structure, and dialogue elaboration, spending 30 minutes on each exercise.
- **Group 2 (Control):** Students in this group used traditional lesson resources for learning. Like the experimental group, they completed three exercises focused on the same three aspects of storytelling, with 30 minutes allocated for each one.

Figure 1.

Exercise example

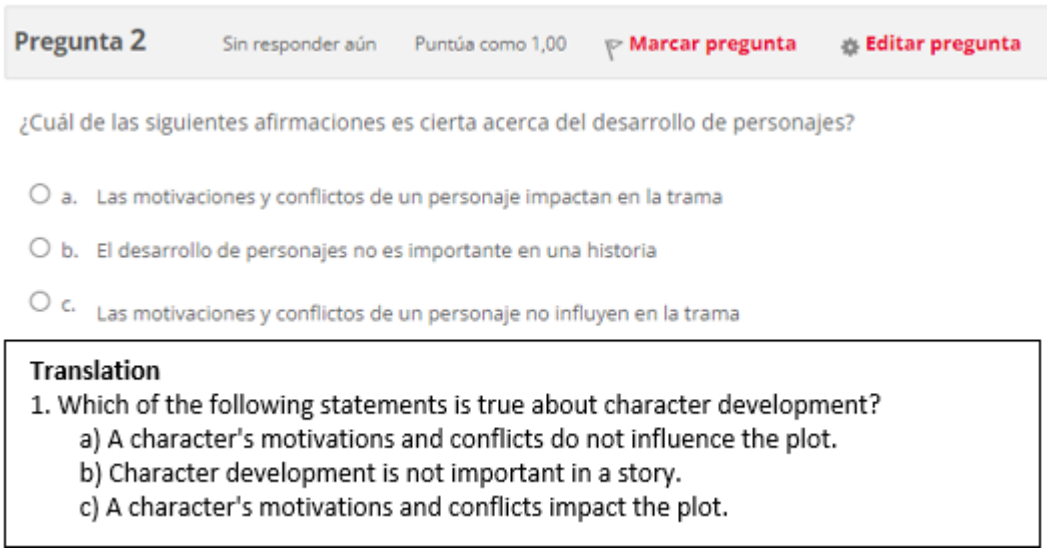
Group 1 (ChatGPT)	**Group 2**
Exercise 1: Character Development (30 minutes) General instructions: Describe an initial situation for your story involving the main characters. You can be as detailed as you want. From this situation, you will develop the characters. Tasks: • Open ChatGPT and provide details about the initial situation. • Use the answers from ChatGPT to create a first version of two main characters and two secondary characters that fit this situation. • Develop additional questions about the characters, based on ChatGPT answers, to delve deeper into their motivations, conflicts, and relationships. • Modify and improve characters according to your preferences and repeat the process with ChatGPT for additional suggestions.	Exercise 1: Character Development (30 minutes) General instructions: Describe an initial situation for your story involving the main characters. Use materials such as articles, YouTube videos, or teacher notes to inspire your character creation. Tasks: • Research and use various resources to obtain information about the initial situation you have raised. • Develop two main characters and two secondary characters based on the research carried out. • Details aspects such as the personality, motivations, conflicts and relationships of the characters. • Adjust and improve characters according to your preferences and the conclusions of your research.

Post-test assessment (20 minutes):

- **Cognitive load questionnaire:** A questionnaire was administered to measure the cognitive load experienced by students, based on the work of Klepsch et al. (2017). This instrument measures the dimensions of intrinsic, extraneous, and germane cognitive load. This questionnaire was implemented in Google Forms.
- **Knowledge test:** A post-intervention knowledge test consisting of 6 questions was conducted to assess knowledge acquisition about digital storytelling following the intervention, using Moodle questionnaire (an example in figure 2)

Figure 2

Test question example



Translation
1. Which of the following statements is true about character development?
   a) A character's motivations and conflicts do not influence the plot.
   b) Character development is not important in a story.
   c) A character's motivations and conflicts impact the plot.

Measurement Instruments

- **Prior and post-intervention knowledge tests:** Designed by the researcher (teacher of the course) to assess the students' level of understanding and knowledge about digital storytelling (focus on character creation, plot structure, and dialogue elaboration) before and after the educational intervention. The type of question was selection, with 3 answer options for each question (see previous section).
- **Cognitive load questionnaire:** Based on the methodology developed by Klepsch et al. (2017), this instrument evaluates the different dimensions of cognitive load (intrinsic, extraneous, and germane) experienced by students during the intervention. The Cronbach's alpha value obtained was approximately 0.749, indicating good internal consistency among the items on the scale. This value suggests that the items are well-correlated and collectively contribute to the construct being measured. A Cronbach's alpha value above 0.7 is generally considered acceptable in social and psychological research, implying that the scale is reliable for measuring the intended construct.

Data Analysis

The data analysis section of this study is focused on examining the differences in digital storytelling knowledge and cognitive load between the experimental group (which used ChatGPT) and the control group (which did not use ChatGPT). To this end, two main types of comparisons were conducted:

1. **Group differences in pre-test.** This comparison aimed to identify any differences between the groups in their pre-test (baseline) (H1).
2. **Intra-group paired differences**: This analysis compared the pre-test and post-test results within each group to determine the extent of change in knowledge about digital storytelling and cognitive load as a result of the intervention (H2).
3. **Group differences in post-test**: This comparison aimed to identify any differences between the groups in the post-test (after intervention) scores (H3).
4. **Group differences in cognitive load**. Additionally, for the post-test, cognitive load between the two groups was also compared to assess the impact of ChatGPT's assistance on easing the mental effort required for digital storytelling (H4).

The statistical analysis was conducted using Jamovi (The jamovi project, 2024), focusing on comparing changes in knowledge levels about digital storytelling and cognitive load between the experimental and control groups. To determine the significance of the observed differences in the results of the knowledge tests and cognitive load questionnaires between both groups, appropriate statistical tests were employed.

Below is a table (table 1) summarizing the results of the Levene's test for equality of variances and tests of normality for both pre-test and post-test scores:

Table 1.
Demographic analysis of total sample respondents.

Homogeneity of Variances Tests		
	statistic	***p***
Pre-test		
Levene's	0.147	0.703
Variance ratio	0.839	0.705
Post-test		
Levene's	0.144	0.706
Variance ratio	0.999	0.999
Tests of Normality		
	statistic	***p***
Pre-test		
Shapiro-Wilk	0.834	< .001
Post-test		
Shapiro-Wilk	0.872	< .001

Source: Authors´ work

The Levene's test results indicate that the variances between the groups are equal for both the pre-test and post-test, as suggested by the non-significant p-values ($p > 0.05$), indicating homogeneity of variances. This is crucial for the validity of subsequent analyses that assume equal variances across groups.

The tests of normality, conducted using the Shapiro-Wilk test, reveal that the data distribution deviates from normality for both pre-test and post-test scores, as indicated by the significant p-values ($p < .001$). Because of the deviation from normality we opting for non-parametric methods for further analysis of the data, given the non-normal distribution of scores

3. Results

3.1. Demographic analysis

In the comprehensive demographic analysis section of our study, we analyzed the characteristics of the 41 participants. The average age of participants in experimental group (EGp) was 36.1 years, while control group (CGp) had an average age of 32.7 years, leading to an overall average age of 34.4 years for the total sample. The analysis, detailed in Table 2, provides insights into the gender composition of our sample, our findings reveal a balanced representation within the sample. The EGp comprised 8 males and 12 females, while the CGp consisted of 7 males and 14 females, showcasing a near-equitable gender ratio within the study population.

Table 2.
Gender analysis of total sample respondents.

	Frecuency		
	EGp	**CGp**	**Total**
Gender			
Male	8	7	21
Female	12	14	20

Source: Authors´ work

3.2. Pre-test results (H1)

The pre-test, conducted to evaluate baseline conditions between the experimental and control groups, demonstrated minimal initial differences ($p = 0.544 > 0.05$). This outcome, as presented in Table 3, confirms the baseline hypothesis (H1) that both groups possess comparable digital storytelling knowledge before the intervention. Such a uniform starting point ensures that any post-test differences can be attributed to the impact of ChatGPT, establishing a solid base for further analysis.

Table 3.
Bias determination between samples through pre-test.

Group	M	SD	Mean difference	p
Grupo 1	0.783	0.248	0.053	0.544
Grupo 2	0.730	0.271		

Source: Authors´ work

3.3. Intra-group Paired Differences (H2)

The intra-group analysis (Table 4), aimed at assessing H2's premise that ChatGPT would significantly enhance digital storytelling skills, shows both experimental and control groups improved their scores. However, the statistical analysis, with p-values exceeding the 0.05 threshold for both groups, leads to the acceptance of H0 for this hypothesis. This outcome suggests that, despite observed improvements in digital storytelling abilities, the differences induced by ChatGPT's use are not statistically significant. Thus, the hypothesis anticipating a marked benefit from ChatGPT in knowledge and skill development in digital storytelling was not supported by the data.

Table 4.
Pre-test and post-test scores of the experimental group and control group.

Experimental Group								
Pre-test		**Post-test**			**Test statistic**			
Mean	**SD**	**Mean**	**SD**	**Mean difference**	**t**	**p**	**Effect Size**	**Decision**
0.783	0.248	0.900	0.166	-0.117	-1.52	0.144	-0.341	Accept H0
Control Group								
Pre-test		Post-test			Test statistic			
Mean	**SD**	**Mean**	**SD**	**Mean difference**	**t**	**p**	**Effect Size**	**Decision**
0.730	0.271	0.817	0.166	-0.0873	-1.39	0.178	-0.304	Accept H0

Source: Authors´ work

3.4. Group differences in post-test (H3)

The post-test comparison (Table 5), aimed at evaluating differences between groups in digital storytelling ability after intervention (H3), revealed a mean score of 0.900 for Group 1 and 0.817 for Group 2, with a mean difference of 0.0825. Despite the observed difference, the p-value of 0.060 suggests that this difference does not reach statistical significance ($p > 0.05$), according to the Mann-Whitney U test. Therefore, while Group 1, which received ChatGPT intervention, showed a higher mean score, this improvement over Group 2, which did not, is

not statistically significant, leading to the acceptance of H0 for H3. This indicates that the impact of ChatGPT on post-intervention digital storytelling skills, while potentially positive, does not demonstrate a clear statistical advantage over traditional methods based on this data set.

Table 5.
Bias determination between samples through post-test.

Group	M	SD	Mean difference	p
Grupo 1	0.900	0.166	0.0825	0.060
Grupo 2	0.817	0.166		

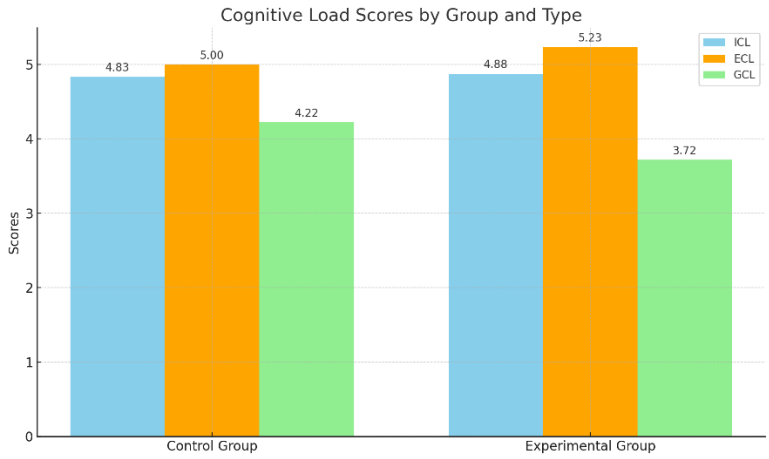
Source: Authors´ work Mann-Whitney U

3.5. Group differences in cognitive load (H4)

The analysis showed in figure 3, considering the three types of cognitive load—namely intrinsic (ICL), extraneous (ECL), and germane cognitive load (GCL)— reveals distinct differences between the experimental group (EGp) and the control group (CGp) following the intervention.

- **Intrinsic Cognitive Load (ICL)** reflects the inherent difficulty of the task. The mean scores slightly increased from the control to the experimental group (4.833 to 4.875), indicating a marginal increase in the perceived complexity and demands of the task. However, the difference (0.042) is minimal, suggesting that both groups perceived the task's inherent difficulty similarly.
- **Extraneous Cognitive Load (ECL)**, associated with the way information or tasks are presented to learners, showed a more noticeable difference. The mean scores increased from 5.000 in the control group to 5.233 in the experimental group, with a difference of 0.233. This suggests that the experimental group, potentially influenced by the intervention, experienced a slightly higher extraneous cognitive load. This could imply that the intervention's design or implementation introduced additional complexity or effort in processing the information.
- **Germane Cognitive Load (GCL)**, which pertains to the mental resources allocated to processing and understanding the task, decreased from the control group to the experimental group (4.222 to 3.717), with a difference of -0.506. This significant decrease suggests that the experimental group found it less taxing to integrate and apply the important information, indicating that the intervention may have effectively facilitated learning by reducing the effort needed to understand and link crucial concepts.

Figure 3.
Group comparison in cognitive load



3.6. Student perceptions (group interview)

The summary of the perceptions of the students in the experimental group about the activities carried out in the experiment with ChatGPT reveals a predominantly positive reception. Participants found the activities interesting, enriching and entertaining, highlighting the interactivity and novelty of using an artificial intelligence tool of this type for the first time. They expressed that they gained practical benefits and appreciated the help that ChatGPT provided on assignments, finding it especially useful for narrative construction and conflict redesign in their writing.

Several students expressed that the activity was an introduction to a digital tool that they found easy to use and applicable to their educational environments. They appreciated how ChatGPT could save time and help the teaching process, as long as they gave specific prompts to generate useful results. The dynamic nature of building scenes and narratives with the help of AI was also seen as a valuable exercise in creativity and knowledge generation.

However, some students faced challenges and described the activity as complicated or difficult, especially for those who were new to the terminology and process of narrative writing with AI. Despite these initial difficulties, the non-evaluative nature of the activity allowed for a pressure-free exploration of the tool, which was appreciated by the participants.

4. Discussion

Demographic Insights and Baseline Equivalence

This study achieves both gender parity and age diversity, thus ensuring a representative sample that reflects the inherent inclusivity of digital storytelling and AI in education. This inclusion is critical to understanding the multifaceted effects of technology on diverse demographic groups. Avello-Martínez et al. (2023) highlight the potential of digital narratives to cultivate inclusive educational environments, indicating that our well-balanced demographics strengthen the generalizability of our findings.

Pre-Test Results and Baseline Knowledge

Establishing baseline equivalence was an important step in our methodology, ensuring that any post-intervention differences could be confidently attributed to the use of ChatGPT. This approach aligns with the methodological rigor required in educational research, as highlighted by (Essel et al., 2024). The minimal initial differences observed reinforce the importance of carefully controlled experimental designs to evaluate the impact of innovative educational technologies. The result, presented in Table 3, confirmed that both groups had similar knowledge about digital storytelling before the intervention. This starting point established a solid foundation for subsequent analyzes.

Intra-Group and Inter-Group Analysis

The observed lack of statistically significant improvement in digital storytelling skills, as determined by intra-group analysis, underscores the complexity of AI's educational impact, particularly regarding ChatGPT. While both experimental and control groups showed improvements, the absence of significant differences underscores the subtlety of AI's role in enhancing educational outcomes. This nuance is echoed in the works of García-Peñalvo (2023) and Muñoz et al. (2024), who argue that AI's effectiveness hinges on thoughtful design and implementation, emphasizing the need for AI to complement rather than replicate traditional pedagogical methods. The call for employing qualitative research methodologies, as Jafari & Keykha (2023) suggest, becomes imperative to capture the depth of students' interactions with AI technologies and their resultant learning experiences. Moreover, the potential for both positive and negative long-term effects of AI on education, highlighted by Pellas (2023), stresses the importance of conducting longitudinal studies. These studies should aim to unravel the enduring impacts of AI on learners' competencies, knowledge acquisition, and attitudes towards learning, offering insights into AI's transformative potential in education. The results from our study, indicating non-significant differences post-ChatGPT intervention, contribute to this ongoing discourse, suggesting that while immediate gains might be subtle, the broader implications and long-term effects of AI in education warrant comprehensive and continued exploration.

Cognitive Load Analysis

Our findings clarify the nuanced interplay between the integration of AI in educational contexts and the management of cognitive load, contributing significantly to the discourse on the effectiveness of digital learning. The subtle increase in intrinsic and extraneous cognitive loads observed following ChatGPT integration underscores the critical need for meticulous instructional design that aligns with human cognitive architecture. This alignment is critical to ensure that the introduction of AI tools into the learning environment does not inadvertently raise cognitive demands to counterproductive levels. In contrast, the notable reduction in relevant cognitive load in the experimental group highlights the potential of AI, such as ChatGPT, to facilitate more efficient learning processes by assisting in the synthesis and application of new knowledge. These findings resonate and expand with the work of Ji et al. (2024), who call for a reevaluation of pedagogical strategies in the digital age, advocating for a balanced approach that leverages the strengths of AI while mitigating its challenges.

5. Conclusions

Our investigation into the effects of ChatGPT on narrative scriptwriting and cognitive load among master's students in educational technology has provided nuanced insights into the

integration of AI in educational settings. Despite achieving gender parity and age diversity, ensuring a representative and inclusive sample, our findings indicate that ChatGPT's impact on enhancing digital storytelling skills is not statistically significant. This outcome suggests a nuanced role of AI in education, emphasizing the importance of design and implementation in leveraging AI tools effectively. The baseline equivalence established between the experimental and control groups allowed for a controlled assessment of ChatGPT's impact, underscoring the methodological rigor essential for evaluating innovative educational technologies.

The intra-group and inter-group analyses further revealed the complexities involved in integrating AI into learning environments. Although improvements were observed in both experimental and control groups, the lack of significant differences post-intervention with ChatGPT highlights the subtlety of AI's role in educational outcomes. This finding aligns with the broader discourse on the need for AI to complement rather than replicate traditional pedagogical methods. Additionally, our cognitive load analysis indicates that while there is a slight increase in intrinsic and extraneous cognitive loads, the significant reduction in germane cognitive load suggests ChatGPT's potential to facilitate more efficient learning processes by aiding in the synthesis and application of new knowledge.

Implications and Future Directions

The contributions of this study to the field of educational technology and AI integration are multifaceted. It highlights the potential of AI tools like ChatGPT to support learning processes, albeit within the context of thoughtful implementation and design. The findings suggest that while immediate gains in digital storytelling skills may not be significant, the reduction in cognitive load offers a promising avenue for enhancing learning efficiency. Future research should focus on longitudinal studies to assess the long-term impact of AI on educational outcomes, exploring how AI tools can be optimized to complement traditional teaching methods and reduce cognitive burdens. Moreover, there is a need for qualitative research to understand better the student experience with AI in learning environments. By continuing to explore these areas, educators and researchers can develop more effective strategies for integrating AI into education, maximizing its potential to enrich learning and teaching while addressing the challenges and limitations of its use.

Presentación del artículo: 30 de enero de 2024
Fecha de aprobación: 7 de abril de 2024
Fecha de publicación: 30 de mayo de 2024

Avello-Martínez, R., Gajderowicz, T., & Gómez-Rodríguez, V.G. (2024). Is ChatGPT helpful for graduate students in acquiring knowledge about digital storytelling and reducing their cognitive load? An experiment. *RED. Revista de Educación a Distancia, 24*(78). http://dx.doi.org/10.6018/red.604621

Authors' statement on the use of LLMs

This article has not used texts from (or generated) from an LLM (ChatGPT or others) for its writing.

Funding

This work has not received any specific grants from funding agencies in the public, commercial, or non-profit sectors.

6. References

Avello-Martínez, R., Gómez-Rodríguez, V., & Rodríguez-Monteagudo, M. (2023). Las narrativas digitales en la formación de postgrado de maestros. Una experiencia en la universidad bolivariana del Ecuador. *Identidad Bolivariana*, *7*(3), Article 3. https://doi.org/10.37611/IB7ol3167-178

Bai, L., Liu, X., & Su, J. (2023). ChatGPT: The cognitive effects on learning and memory. *Brain-X*, *1*(3), e30. https://doi.org/10.1002/brx2.30

Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). *Computers and Education: Artificial Intelligence*, *6*, 100198. https://doi.org/10.1016/j.caeai.2023.100198

Fang, X., Ng, D. T. K., Leung, J. K. L., & Chu, S. K. W. (2023). A systematic review of artificial intelligence technologies used for story writing. *Education and Information Technologies*, *28*(11), 14361-14397. https://doi.org/10.1007/s10639-023-11741-5

García-Peñalvo, F. J. (2023). La percepción de la Inteligencia Artificial en contextos educativos tras el lanzamiento de ChatGPT: Disrupción o pánico. *Education in the Knowledge Society (EKS)*, *24*, e31279-e31279. https://doi.org/10.14201/eks.31279

Hurtado-Mazeyra, A., Alejandro-Oviedo, O. M., Núñez-Pacheco, R., & Almenara, J. C. (2023). El Digital Storytelling en la modalidad 2D y con realidad aumentada para el desarrollo de la creatividad en la educación infantil. *Revista de Educación a Distancia (RED)*, *23*(73), Article 73. https://doi.org/10.6018/red.536641

Jafari, F., & Keykha, A. (2023). Identifying the opportunities and challenges of artificial intelligence in higher education: A qualitative study. *Journal of Applied Research in Higher Education*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/JARHE-09-2023-0426

Ji, Y., Zou, X., Li, T., & Zhan, Z. (2024). The effectiveness of ChatGPT on pre-service teachers' STEM teaching literacy, learning performance, and cognitive load in a teacher training course. *Proceedings of the 2023 6th International Conference on Educational Technology Management*, 16-22. https://doi.org/10.1145/3637907.3637948

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.01997

Pellas, N. (2023). The Effects of Generative AI Platforms on Undergraduates' Narrative Intelligence and Writing Self-Efficacy. *Education Sciences*, *13*(11), Article 11. https://doi.org/10.3390/educsci13111155

Rodriguez, C. L., Garcia-Jimenez, M., Masso-Guijarro, B., & Cruz-Gonzalez, C. (2021). Digital Storytelling in Education: A Systematic Review of the Literature. *Review of European Studies*, *13*, 13.

The jamovi project. (2024). *Jamovi (Version 2.3)[Computer Software]*. https://www.jamovi.org/about.html

van Gog, T., Paas, F., & Sweller, J. (2010). Cognitive Load Theory: Advances in Research on Worked Examples, Animations, and Cognitive Load Measurement. *Educational Psychology Review*, *22*(4), 375-378. https://doi.org/10.1007/s10648-010-9145-4

Vandewaetere, M., & Clarebout, G. (2013). Cognitive Load of Learner Control: Extraneous or Germane Load? *Education Research International*, *2013*, e902809. https://doi.org/10.1155/2013/902809

Wu, J., & Chen, D.-T. V. (2020). A systematic review of educational digital storytelling. *Computers & Education*, *147*, 103786. https://doi.org/10.1016/j.compedu.2019.103786