







# Writing, creativity, and artificial intelligence. ChatGPT in the university context

Escritura, creatividad e inteligencia artificial.

ChatGPT en el contexto universitario

-  Dr. María-Isabel de Vicente-Yagüe-Jara. Senior Lecturer, Department of Didactics of Language and Literature, University of Murcia (Spain) (isabelvyague@um.es) (<https://orcid.org/0000-0002-2496-2971>)
-  Dr. Olivia López-Martínez. Senior Lecturer, Department of Developmental and Educational Psychology, University of Murcia (Spain) (olivia@um.es) (<https://orcid.org/0000-0002-9819-8005>)
-  Dr. Verónica Navarro-Navarro. Lecturer, Department of History of Art & Fine Arts, University of Salamanca (Spain) (veronicanavarro@usal.es) (<https://orcid.org/0000-0003-3439-9278>)
-  Dr. Francisco Cuéllar-Santiago. Lecturer, Art Department, Miguel Hernández University of Elche (Spain) (fcuellar@umh.es) (<https://orcid.org/0000-0003-0523-6509>)

## ABSTRACT

The main objective of the research is to study the creative potential of Artificial Intelligence (AI) for writing skills in an educational context. The research aims to provide evidence on the use of AI and contribute to its integration in the classroom as a support for the teaching-learning process. Two types of research designs were established: a descriptive and comparative non-experimental quantitative research, and a quasi-experimental pretest-posttest study. The sample consisted of 20 AI systems and 193 university students who were given Games 2 and 3 of the Spanish PIC-A test ("Creative Imagination Test for Adults"). The students repeated the games, assisted by ChatGPT, to compare the possible improvement of their productions. The findings reveal statistically significant differences between the AIs and the students in the indicators of fluency, flexibility, and narrative originality in Game 2. Furthermore, significant differences are found between students' pre-test and post-test scores in fluency, flexibility, and narrative originality in Game 2 and in fluency in Game 3. Finally, the assistance provided by AI in writing tasks and verbal creativity is highlighted, and this should be considered in language teaching; in any case, AI cannot replace human intelligence and creativity.

## RESUMEN

La investigación persigue estudiar las posibilidades creativas de los sistemas de Inteligencia Artificial (IA) para el desarrollo de la escritura en el contexto educativo. Se persigue aportar evidencias en el uso de la IA y contribuir al conocimiento de su integración en las aulas como apoyo al proceso de enseñanza-aprendizaje. Se establecen dos tipos de diseño: una investigación de corte cuantitativo no experimental de tipo descriptivo y comparativo, y un estudio cuasi-experimental de tipo pretest-postest. La muestra estuvo compuesta por 20 sistemas de IA y 193 estudiantes universitarios, a los cuales se les aplicaron los juegos 2 y 3 del test español PIC-A («Prueba de Imaginación Creativa para Adultos»). El alumnado repitió los juegos con ayuda de ChatGPT, con el fin de comparar la posible mejora de sus producciones. Los resultados destacan la existencia de diferencias estadísticamente significativas entre las IA y el alumnado en los indicadores de fluidez, flexibilidad y originalidad narrativa del juego 2. Además, se encuentran diferencias significativas entre las puntuaciones del pretest y postest del alumnado en fluidez, flexibilidad y originalidad narrativa del juego 2, así como en fluidez del juego 3. Finalmente, se pone de manifiesto la ayuda que la IA proporciona en tareas de escritura y creatividad verbal, lo que debería ser tenido en cuenta en la enseñanza de lenguas; en cualquier caso, la IA no puede reemplazar a la inteligencia y la creatividad humana.

## KEYWORDS | PALABRAS CLAVE

Artificial intelligence, writing, language teaching, verbal creativity, ChatGPT, Large Language Models.  
Inteligencia artificial, escritura, enseñanza de lenguas, creatividad verbal, ChatGPT, modelos de lenguaje extensivos.



## 1. Introduction

The rapid development of Artificial Intelligence (AI) is a reality that involves multiple opportunities, risks, and challenges in the field of education, which have so far outpaced policies and legislative frameworks. In this regard, UNESCO (2019) committed to harnessing the potential of AI technologies in order to advance towards Sustainable Development Goal 4 (ensure inclusive and equitable quality education and promote lifelong learning opportunities for all) and achieve the Education 2030 Agenda. AI originally emerged as a tool to simulate and mechanise human thought processes (Turing, 1950). Today, it has become the basic grammar of our century that aims to develop citizens' AI literacy and competencies (UNESCO, 2021). Moreover, UNESCO aims to achieve a human-centred approach, based on principles of inclusion and equity, in order not to widen technological gaps and to ensure "AI for all" in terms of innovation and knowledge. In this sense, one of the most important challenges is to ensure that AI is designed and used in an ethical and responsible way, in order to avoid misuse of technology or the increase of existing inequalities in society (UNESCO, 2022a).

Recently, the OECD (2021) considers the ways in which smart technologies are changing classroom education and the management of educational organisations and systems. Nowadays, there are already various ways in which AI can be involved in teaching, from OER (Open Educational Resources) content recommendation, student emotion detection, intelligent tutoring systems, AI-powered teaching assistants, to automatic exam marking and automatic forum monitoring (Flores-Vivar & García-Peñalvo, 2023).

There are several AI systems that are used to generate text automatically ("Large Language Model" or LLM). Their origin dates back to 2017, when the architecture of Transformer Models (Vaswani et al., 2017), neural networks that can learn through attention mechanisms, was presented. Experiments on two machine translation tasks showed that these models are highly efficient in terms of output quality and lower training. This model can learn context through sequential data, which is considered by experts to be the beginning of LLMs. Subsequently, in 2018, Google launched a research project based on Natural Language Processing (NLP), whose technology sought the interaction between human and computer dialects, bringing together the disciplines of Applied Linguistics, Computer Science, and AI. Google presented BERT ("Bidirectional Encoder Representations from Transformers"); a state-of-the-art NLP system that would help search work through its well-known search engine. Thus, LLMs are deep learning algorithms that can recognise, summarise, translate, predict, and generate text or other content based on knowledge acquired from massive data sets. This learning is unsupervised, as it feeds a given amount of data to an AI without explicit instructions on what to do with it. Among the various applications of this large language model (Lee, 2023) are ChatGPT and other systems, which have been trained to answer questions or follow specific writing instructions.

These AI systems can write in a particular tone (humorous, familiar, professional, witty, friendly), rewrite or paraphrase sections of a given text, write after a title, or write in Shakespearean style. They represent an important challenge for language teaching and, in particular, for work on written expression and the development of creative writing. Creativity is a capacity of the human mind (Csikszentmihalyi, 1998; Guilford, 1950; Sternberg, 1999), but, as Boden (2004) pointed out, computers and creativity can be interesting partners in two ways: for understanding human creativity and for producing computational creativity. This author explains creativity from a scientific approach that uses computational concepts from the field of AI, concepts that allow us to create and test hypotheses about the structures and processes that may be involved in human thought.

With respect to computational creativity, it should be noted that humans play a fundamental role in programming, choosing models, and fine-tuning AI systems. However, if we consider Boden's (2004) three types of creativity (combinatorial, exploratory, and transformational), computers can generate ideas that at least appear to be creative. Moreover, he argues that computers can produce new ideas and help people to do so; both their failures and their successes make it possible to think more clearly about the creative power of human beings. Miller (2019) also questions whether AI can be creative, based on a review of various applications in the fields of visual arts, music, text, and musical theatre. He identifies the essential factors for the creative process, from the need for introspection to the ability to discover the key problem, and concludes that computers can be as creative as humans. On the other

hand, Ward (2020) suggests that computational creativity is not (and does not need to be) equivalent to human creativity but could be different and even contribute to new processes and outcomes that could be considered creative; machines do not need to be more like humans, but humans could recognise the creative capabilities inherent in the machine. This is therefore the interesting direction that the new era of AI could take for the development of teaching and learning processes.

In this sense, language teaching should not avoid AI systems, considering the growing expansion of digital environments for writing and the significant presence of such environments in educational contexts. Writing, AI, and creativity become fundamental aims for new models of language teaching. Moreover, creativity cannot be separated from culture and the use or learning of a language (Argondizzo, 2012). Along these lines, creative thinking has taken on an important educational dimension as it has been included as a new assessable competence in PISA (OECD, 2019); this programme assesses the area of creative expression (where the domain of both written and visual expression is located), in addition to the area of knowledge creation and creative problem-solving finding creativity in artistic and inventive practices, this is an effective problem-solving in all types of challenging situations (Rodrigo-Martín et al., 2022), which is why it is advisable to develop it in the educational environment. Moreover, creative thinking is identified as one of the prerequisite competencies which should be promoted by Member States for AI education, according to UNESCO (2022a).

Considering all the above, chatbots and the communicative interaction enabled by AI systems pose a challenge to the teaching practice of language teachers. Several studies have shown that the integration of AI improves the quality and effectiveness of foreign language teaching, as it favours an individualised and cooperative learning style (Sun et al., 2021; Yanhua, 2020). In this sense, given the variety of tools, programmes, and resources available to learners, it is crucial to study their possibilities, limitations, and ways of use, both for the benefit of adequate acquisition of the necessary competencies and for ethical use. Moreover, in the educational environment, AI systems must be subject to strict requirements for monitoring, assessing learners' skills or predicting their behaviour. AI must support the learning process without reducing cognitive abilities. Furthermore, the information collected by the learner in their interactions with AI systems shall not be subject to illicit use, misappropriation, or criminal exploitation (UNESCO, 2022a).

Several initiatives have been addressing its potential educational use. For example, UNESCO (n.d.) is developing the portal "Teaching AI for K-12" (students aged 5-18), which includes resources to be used by teachers. In addition, in order to ensure that AI writing support tools use correct Spanish, the Royal Spanish Academy (RAE, n.d.) has led the project "Lengua Española e Inteligencia Artificial" (LEIA in Spanish) [Spanish Language and Artificial Intelligence]. It has also developed the MarIA project (Gutiérrez-Fandiño et al., 2022), a family of large language models in Spanish, made available to industry and research, which uses the RoBERTa-base, RoBERTa-large, GPT2 and GPT2-large models for their construction. MarIA has been created to solve the existing conflict with languages in GPT, as most LLM developers have been built in English. In addition, AI is already being used as a co-worker in some companies, according to Montero (2023), who recently pointed out that OpenAI chat has been included in a design project he is working on. In education, AI is also beginning to be experimented with, as Hendrickson, a lecturer in Digital Media at the University of Leeds, has her students use both ChatGPT and other AI language models in their writing assignments to improve their academic writing, while also testing how the technology works, enabling students to critically reflect on and question this text-generating resource (Renbarger, 2023). Similarly, the present research aims to approach the use of AI in teaching classrooms, as a companion and support, in order to integrate its use and check what kind of support this can provide. In the absence of solid and proven evidence on the effectiveness and impact of AI on students' academic achievement, research is needed in this area. Research needs to be carried out in a future that is yet to be defined, as AI is a very powerful technology and the challenge is to "discover ways to use it with meaningfulness and awareness" (Selwyn et al., 2022: 143).

Therefore, the general objective of the present research is aimed at studying the creative possibilities of AI systems for the development of writing in the educational context. This general objective is articulated in the following specific objectives:

- To find out the level of creativity indicators (fluency, flexibility, and narrative originality) of AI systems and students, based on the sample of the Creative Imagination Test for Adults (PIC-A) by Artola et al. (2012).
- To compare the creativity of the AI systems with that of the students, based on the indicators of fluency, flexibility, and narrative originality.
- To compare the scores of the creativity indicators obtained by the students in the two phases of application of the tests, the first without any type of aid and the second using an AI system (ChatGPT).

## 2. Materials and methods

### 2.1. Design

Aiming at addressing the objectives formulated, two types of research design were established. Firstly, a descriptive and comparative non-experimental or ex post-facto quantitative research study was carried out, in which the level of creativity achieved by both the AI systems and the students was analysed, and the extent of the relationship between the creativity of the AIs and the students was studied.

Secondly, a quasi-experimental pretest-posttest study was carried out, where the creativity indicators obtained by the students, both before and after a learning intervention with the help of ChatGPT, were evaluated and compared.

### 2.2. Sample and participants

The sample consisted of 20 Large Language Model (LLM) systems with OpenAI/GPT-3 technology (except Dupla.ai, a Chrome extension), which have the function of automatically generating text from a given instruction (Table 1). In the selection of AI systems, the following inclusion criteria were taken into account: open access (free of charge or “Freemium”) and online execution without prior download. These criteria are explained by the intended use of AI in an educational context.

Similarly, 193 students (153 women and 40 men) aged 18 to 50 years old, enrolled in the Degree in Primary Education and the Master’s Degree in Research and Innovation in Early Childhood Education and Primary Education at the University of Murcia, the Degree in Fine Arts at the University of Salamanca and the double Degree in Audio-visual Communication and Journalism at the Miguel Hernández University of Elche participated in the study. A non-probabilistic, purposive, convenience sampling procedure was used.

Name	Website	Organisation	Access
ChatGPT	<a href="https://chat.openai.com">https://chat.openai.com</a>	OpenAI	Free
Copy.ai	<a href="https://www.copy.ai">https://www.copy.ai</a>	CopyAI, Inc	Freemium
Rytr	<a href="https://rytr.me">https://rytr.me</a>	Rytr LLC	Freemium
Peppertype	<a href="https://www.peppertype.ai">https://www.peppertype.ai</a>	Pepper Content Pvt. Ltd	Paid/Freemium
Copysmith	<a href="https://copysmith.ai">https://copysmith.ai</a>	Copysmith Artificial IntelligenceInc	Paid/Freemium
Jasper	<a href="https://www.jasper.ai">https://www.jasper.ai</a>	Jasper AI, Inc.	Paid/Freemium
Writesonic	<a href="https://writesonic.com">https://writesonic.com</a>	Writesonic, Inc.	Freemium
Youchat	<a href="https://youchat.com">https://youchat.com</a>	YouChat Inc	Free
Anyword	<a href="https://anyword.com">https://anyword.com</a>	Keywee Inc. (d.b.a. Anyword)	Freemium
ASKtoAI	<a href="https://asktoai.com/es/">https://asktoai.com/es/</a>	MakeYourCloud LTD	Paid/Freemium
Contentbot	<a href="https://contentbot.ai">https://contentbot.ai</a>	ContentBot.ai	Paid/Freemium
Copymatic	<a href="https://copymatic.ai/app/">https://copymatic.ai/app/</a>	Copymatic Inc	Paid/Freemium
Dupla	<a href="https://app.dupla.ai">https://app.dupla.ai</a>	duplaAI, Inc	Chrome extension
Magic Write (Canva)	<a href="https://www.canva.com">https://www.canva.com</a>	Canva Pty Ltd	Freemium
Neuro-Flash	<a href="https://app.neuro-flash.com">https://app.neuro-flash.com</a>	Neuroflash GmbH	Paid/Freemium
Nichesss	<a href="https://nichesss.com">https://nichesss.com</a>	Malcolmytson	Freemium
ReAI	<a href="https://ai.devinay.in/?ref=producthunt">https://ai.devinay.in/?ref=producthunt</a>	Regem Enterprises	Paid/Freemium
Simplified	<a href="https://app.simplified.com/ai-writer">https://app.simplified.com/ai-writer</a>	TLDR Technologies, Inc.	Freemium
Unbounce	<a href="https://unbounce.com">https://unbounce.com</a>	Unbounce Marketing Solutions Inc.	Paid/Freemium
Wordtune	<a href="https://www.wordtune.com">https://www.wordtune.com</a>	AI21 Labs	Freemium

### 2.3. Tools

The instrument used was the Test of Creative Imagination for Adults (PIC-A) by Artola et al. (2012). Of the four games that make up the test, the second and third were used, as these are the ones that focus exclusively on verbal text. The instructions for the games are as follows:

- Game 2. Make a list of all the things a rubber tube could be used for. Think of interesting and original things. Write down all the uses you could have for it, even if they are imagined. You can use any number and size of tubes you want. Example: “As a water pipe”.
- Game 3. Imagine and answer what you think would happen if this sentence were true: What would happen if people never stopped growing? Example: “Stretchy clothes would sell much more”.

The indicators of creativity, assessed in both the second and third games, are as follows:

- Fluency: the ability to produce a large number of ideas. A high fluency score corresponds to being able to make a large number of associations with a stimulus.
- Flexibility: the ability to produce a wide variety of responses, related to different domains. A high flexibility score relates to the ability to search for solutions using a variety of alternatives and the ability to change perspective.
- Narrative originality: the ability to produce ideas that are far from the obvious or established. A response is considered original when its frequency of occurrence is very low.

Finally, reliability was calculated using Cronbach’s alpha, which indicated that both game 2 ( $\alpha = .803$ ) and game 3 ( $\alpha = .853$ ) showed high internal consistency.

#### 2.4. Procedure

Firstly, each of the AI systems was asked to perform the two sets of the PIC-A. The instruction for each test was typed into the chats and text boxes of the systems and the response provided was collected. Given the speed with which the systems generate text, instead of using the time set by the instrument, the information was requested on three continuous occasions, sending the same instruction each time. Subsequently, the automatic text generated was corrected, based on the indicators of fluency, flexibility, and narrative originality, according to the correction notebook established in the instrument (Artola et al., 2012). It should be noted that bearing in mind that the AI learns from previous interactions, the two PIC-A games were performed as the first action once the system was accessed, to ensure that no circumstance could have influenced its response.

As far as the research with the students is concerned, one-hour sessions were used in each of the groups. Firstly, in the pre-test phase, the students played the two creativity games, following the instructions and duration provided in the manual (10 minutes each). Secondly, the students registered in ChatGPT (the AI selected for its significant diffusion compared to others designed previously, its media impact, and the current debate on its use in the educational field), its functioning was explained, and they carried out test queries. Thirdly, the students performed the games again (post-test phase), this time with the help of the answers provided by ChatGPT; it was explained to them that they had to improve the content they had written when they took the test for the first time, so it was not a question of taking the tests again without taking into account what they had done previously. The learners could ask the chat what they considered in order to improve their text, either by asking for the test instruction literally or by asking other questions that could be aimed at improving their written productions.

Participants were informed about the confidentiality of the data and research objectives, following the ethical standards indicated in The Code of Good Practice on Research of the University of Murcia (2022), and any doubts that arose at the time of application were addressed.

#### 2.5. Data analysis

The indicator scores for creativity were calculated according to the correction instructions and category system of Artola et al. (2012). The direct scores are calculated from the following indications: a) Fluency is obtained by counting the total number of different answers given; b) Flexibility refers to the number of different categories (according to the established scale) in which at least one answer is classified; c) Narrative originality is the result of multiplying the number of answers in each category by the coefficient given to that category and their total sum. The determination of high or low creativity is made on the basis of the table of averages of the indicators in the different games, provided by the instrument manual (Artola et al., 2012: 86).

Once the correction had been made, the descriptive statistics of the different variables were analysed in order to find out the level of creativity of the AI systems and of the students, according to the first specific objective. The Kolmogorov-Smirnov normality test was then applied to determine the relevance of the use of parametric tests. This test showed that fluency ( $p < .05$ ), flexibility ( $p < .05$ ), and narrative originality ( $p < .05$ ) did not follow a normal distribution. Thus, in order to compare the creativity of the AI systems and the students, as pursued by the second specific objective, the Mann-Whitney U test was performed. Finally, to answer the third specific objective, the Wilcoxon test was performed.

### 3. Results

The results are presented below according to the specific objectives formulated. The first specific objective focused on finding out the level of the creativity indicators of the AI systems and the students. A descriptive statistical analysis was carried out of the scores obtained after correcting the games played by the AI systems and the students in terms of fluency, flexibility, and narrative originality. Firstly, with respect to the AI systems (Table 2), it was observed that the means obtained in game 2 are higher than the means indicated in the instrument manual for the indicators of fluency ( $M=15.54$ ), flexibility ( $M=9.75$ ) and narrative originality ( $M=11.15$ ). However, in game 3, the means are lower than those given in the manual for fluency ( $M=13.78$ ), flexibility ( $M=8.75$ ) and narrative originality ( $M=8.12$ ) (Artola et al., 2012: 86).

In game 2, none of the AIs falls below the averages for fluency and narrative originality indicated in the manual. In fluency, ASKtoAI (DS-direct score: 81), Youchat (DS: 65), Writesonic (PD: 41) and Dupla (PD: 40) stand out. In narrative originality, ASKtoAI (DS: 87), ChatGPT (DS: 85), Copy.AI (DS: 60), Unbounce (DS: 49) and Anyword (DS: 47) stand out. In flexibility, not all AIs exceed the manual average, although Dupla (DS: 29), Canva's Magic Write (DS: 19) and Anyword (DS: 18) should be mentioned.

In game 3, the AIs with high scores in fluency are Neuroflash (DS: 30), ASKtoAI (DS: 20) and Jasper (DS: 20). In flexibility, Neuroflash (DS: 20), Writesonic (DS: 11), ChatGPT (DS: 10), ASKtoAI (DS: 10) and Nichesss (DS: 10) stand out. In narrative originality, ChatGPT (DS: 25), ASKtoAI (DS: 22) and Neuroflash (DS: 18) scored highest.

**Table 2. Descriptive statistics for indicators of creativity in AI systems**

	G2 Fluency	G2 Flexibility	G2 Narrative Originality		G3 Fluency	G3 Flexibility	G3 Narrative Originality	
M	29,3	12,35	37,4		10,65	6,90	7,80	
SD	16,9	5,33	20,5		7,44	4,34	7,82	
Minimum	14	5	12		1	1	0	
Maximum	81	29	87		30	20	25	

Secondly, after performing the descriptive statistical analysis of the students' scores (Table 3), it was discovered that the sample of the present study presented values below the means of fluency and flexibility of the two games, as well as below the means of narrative originality of game 2, as indicated in the manual of the Creative Imagination Test for Adults (Artola et al., 2012: 86).

**Table 3. Descriptive statistics of the indicators of creativity in students**

	G2 Fluency		G2 Flexibility		G2 Narrative Originality		G3 Fluency		G3 Flexibility		G3 Narrative Originality	
	M	W	M	W	M	W	M	W	M	W	M	W
M	10,8	9,05	7,61	7,38	11,2	10,2	9,87	8,20	6,95	6,53	9,84	7,47
	10,47		7,56		11,02		9,52		6,87		9,35	
SD	5,76	5,59	3,60	3,81	8,25	8,71	5,66	5,09	3,33	3,97	8,89	6,61
Minimum	2	1	2	1	0	0	1	1	1	1	0	0
Maximum	31	27	18	17	43	39	34	22	22	16	52	22

However, the mean of narrative originality of game 3 was above the mean indicated by the instrument referred to. If we look at the scores according to gender, women obtained a better mean in the narrative originality of both games, compared to that presented in the instrument's manual. In general, women showed better mean scores than men in all the creativity indicators in the sample of the present study.

For the second specific objective, in order to compare the creativity of the AI systems and the students, the Mann-Whitney U test was performed, which shows that there are statistically significant differences in the indicators of fluency, flexibility and originality in game 2 (Table 4).

**Table 4. Creativity of AI systems and learners**

Creativity indicators	Human Intelligence		Artificial Intelligence		U	Z	p	Effect Size
	M	SD	M	SD				
G2 Fluency	10,47	5,756	29,30	16,927	228,000	-6,499	<,001	,445
G2 Flexibility	7,56	3,637	12,35	5,334	790,000	-4,363	<,001	,299
G2 Narrative Originality	11,02	8,337	37,40	20,451	246,500	-6,422	<,001	,440
G3 Fluency	9,52	5,576	10,65	7,443	1798,000	-,504	,614	,034
G3 Flexibility	6,87	3,467	6,90	4,340	1913,000	-,065	,948	,004
G3 Narrative Originality	9,35	8,509	7,80	7,818	1644,500	-1,090	,276	,075

Moreover, no statistically significant differences were found in the indicators of game 3, indicating similar levels between the students (human intelligence) and the AI. However, it should be noted that students show a higher mean on the narrative originality indicator in game 3. Also, fluency and narrative originality in game 2 have moderate effect sizes, while the other indicators in game 2 and game 3 have low effect sizes.

Following this, to address the third specific objective, aimed at comparing the scores of the creativity indicators obtained by the students in the two phases of application of the tests, the Wilcoxon test was performed, which revealed the existence of statistically significant differences between the pre-test and post-test scores in fluency, flexibility, and originality in game 2, as well as in fluency in game 3 (Table 5). On the other hand, fluency, and narrative originality in game 2 show a moderate effect size, while both the flexibility indicator in game 2 and the three indicators in game 3 show a low effect size.

**Table 5. Students' pretest and posttest**

Creativity indicators	Pretest		Posttest		Z	p	Effect size
	M	SD	M	SD			
G2 Fluency	10,47	5,756	13,08	7,554	-5,666 <sup>b</sup>	<,001	,408
G2 Flexibility	7,56	3,637	8,33	3,958	-3,036 <sup>b</sup>	,002	,219
G2 Narrative originality	11,02	8,337	14,45	10,816	-5,583 <sup>b</sup>	<,001	,402
G3 Fluency	9,52	5,576	10,05	5,774	-2,236 <sup>b</sup>	,025	,161
G3 Flexibility	6,87	3,467	7,15	3,782	-1,738 <sup>a</sup>	,082	,125
G3 Narrative originality	9,35	8,509	9,31	8,590	-1,596 <sup>b</sup>	,110	,115

#### 4. Discussion and conclusions

The present research aimed to approach the study of the creative possibilities of AI systems for the development of writing in the educational context. We have managed to provide didactic evidence on the use of AI and to contribute to the knowledge of its integration in classrooms as educational support. The role that AI systems can assume in the new technological framework, where the writing-creativity-AI trinomial can begin to be contemplated in the new models of language teaching, has been verified.

Firstly, it is noticeable that the means obtained by the AIs in the three indicators of game 2 are higher than the means indicated in the instrument manual, while, on the contrary, the means achieved in the indicators of game 3 are lower. This situation is related to the type of instruction requested in each of the games and highlights the difference between AI systems and humans. While game 2 explores the possible uses of a real object, the statement in game 3 starts from an implausible situation and aims to assess a fanciful aspect of imagination, a form of thinking that is fundamental to creative behaviour. Uncreative subjects are unable to conceive of such possibilities in this context and may clash when playing this game (Artola et al., 2012), as has happened to AI systems. Moreover, play involves insight into the experience, and while certain consequences are easy to discover, others require deepening of the subject matter and problems (Guilford, 1950). Csikszentmihalyi (1998) already argued that creativity does not happen in an isolated mind, but in the interaction of the person with a socio-cultural context, with established symbolic rules and a field of experts who recognise innovation. Also, taking into account that fantasy (compared to the invention, creativity, and imagination) is the activity that is presented as the freest, creativity being the sum of fantasy and invention (Munari, 2018), it is observed that AI responds more effectively to the invention, seeking practical solutions. In relation to these results, it is striking that there is an AI (pymetrics.ai), as described by Sadin (2018), dedicated to selecting personnel for companies, based on their creativity, adaptability, and flexibility. In this vein, Miller (2019) argues that to be truly creative, machines will need

to enter the world. In relation to the students, as has been noted in other studies (Donolo & Elisondo, 2007), females have a higher level than males in the sample in all creativity indicators, as well as a higher level of narrative originality in both games in relation to the scale of averages in the instrument manual. In general, it can be stated that gender differences in creativity depend on social and cultural factors (Aleksi et al., 2016; Hora et al., 2022). In addition, after comparing the creativity of the AI systems with that of the students, the AI scores on the indicators in game 2 significantly outperformed the students' scores, while in game 3 the creativity indicators showed similar levels in both groups. This result is consistent with the above, as AI scores were higher in game 2. AIs have a large amount of information, more than a person can handle and, as Goleman et al. (2016) stated, having accurate and abundant data is essential in the creative process, since the greater the knowledge of the details of a problem, the greater the probability of finding a solution. With respect to game 3, the students outperform, although not significantly, the AI in narrative originality. Humans do not possess the rigidity that AI systems have exhibited, not being able to distance themselves from the statement in game 3 and find answers that could be pigeonholed into surprising, new, non-obvious, and valuable categories (Boden, 2004; 2018).

Furthermore, the significant improvement in the student's written productions after the use of ChatGPT has been highlighted, specifically in all the indicators of game 2 and the fluency of game 3. The potential of human-AI collaboration in the writing process is highlighted, as "AI is a means, not an end" (Breton, 2021) and it has been a tool that has worked as an assistant to the students in game 2, which is aligned with the productive possibilities of AI technology, whereas the flexibility and originality of game 3 has not, as it has already been shown that the AIs have not been able to imagine the effects of a fantasy situation. In addition to the explanations already offered, this result may also be due to what is called the "framing problem", which refers to the fact that systems programmed for a specific purpose perform better, although they are not able to successfully perform another activity for which they are not programmed (Bonami et al., 2020).

In light of these findings, writing with computational assistance suggests a rethinking of how writing and creativity are conceived. There are questions and concerns with academic honesty and plagiarism, as reflected in recent scientific (Else, 2023; O'Connor and ChatGPT, 2023) and mainstream (Sánchez, 2023; Vázquez, 2023) publications; and indeed, in parallel to the growth of AIs, methods are being developed to detect texts written by them, such as "Classifier" (OpenAI, 2023). In any case, banning AI systems in education (Peirón, 2023) would be a losing battle, so instead of pretending that AI does not exist, it is time to train students to work with it; in the same way that tools such as ChatGPT are being used in science and several researchers confirm their use (Hutson, 2022). Teachers should therefore reflect on the skills they teach and how AI could help students in generating ideas and developing their creativity when dealing with writing tasks.

On the other hand, having proven the AI's assistance in terms of verbal creativity, it must be pointed out that AI cannot replace human intelligence and creativity. Montero (2023) pointed out that AI lacks judgment and stated that creativity is not having ideas but knowing when not to have more. The machine lacks context and does not know what is good to be valued. Therefore, the active participation of students in their learning process should be encouraged and they should be encouraged to act decisively, judiciously, and responsibly. Collaborative human-machine intelligence is required, as highlighted by the Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado [National Institute of Educational Technologies and Teacher Training] (INTEF, 2022), noting the complex relationship between people and AI, as there are specific functions that can only be performed by humans and for which they must be trained (Holmes et al., 2019). Along these lines, Fyfe's study (2022), like the present research, asked university students to conduct an essay using AI. Initially, applying AI to writing may have seemed like a shortcut, but the artificially generated text was difficult to control, deviated from the topic and had to be revised in its different samples, which did not allow for the automatic integration of information in an extensive genre such as the essay.

On the basis of this study, further research in the field of writing is of interest, by extending the study of creativity with the help of AI by means of narrative composition tasks; along these lines, books on the use of ChatGPT to write and structure long texts are beginning to appear on the market (Gade, 2023). In



addition, it is useful to broaden the human sample of participants to include other degrees or professional areas. Similarly, it is possible to envisage future research that delves deeper into the different domains of creativity, in line with the rapid advancement of AI in the creation of a wide variety of products. Thus, the present research can be completed with other studies focused on the creation of images or musical sequences using AI. In short, there is a need for more in-depth studies on the use and exploitation of AI and, specifically, ChatGPT in an educational context, as textbook publishers themselves are beginning to incorporate these resources into their content platforms, as is the case of Edelvives' intelligent assistant that integrates ChatGPT (Edelvives, 2023). Lastly, the use of AI should be aimed at empowering teachers and not replacing or displacing them (UNESCO, 2022b). An AI Literacy Plan is required to train teachers in both technical skills and ethical-philosophical debates (Flores-Vivar & García-Peñalvo, 2023; UNESCO, 2022a). Similarly, changes in educational practice in the coming years will be determined by AI developments and will need to rely on research in full collaboration with teachers, educational leaders, and learners to ensure that appropriate educational policies are put in place (OECD, 2021).

### Authors' Contribution

Idea, M.I.V.Y.J., O.L.M., V.N.N., F.C.S.; Literature review (state of the art), M.I.V.Y.J., F.C.S.; Methodology, M.I.V.Y.J., O.L.M.; Data Analysis, M.I.V.Y.J., O.L.M., V.N.N.; Results, M.I.V.Y.J.; Discussion and conclusions, M.I.V.Y.J., O.L.M., V.N.N.; Drafting (original draft), M.I.V.Y.J.; Final revisions, M.I.V.Y.J., O.L.M., V.N.N., F.C.S.; Project design and funding, O.L.M., M.I.V.Y.J.

### Funding Agency

The study was carried out within the framework of the R&D&I project "Creatividad verbal como dominio: diseño, validación y aplicación de una prueba fundamentada en la comprensión de textos" ["Verbal creativity as a domain: design, validation and application of a test based on text comprehension"] (PID2020-113731GB-I00), funded by MCIN/AEI/10.13039/501100011033. Modalities "Research Challenges" and "Knowledge Generation" (without ERDF funding). Interdisciplinary Research Group: Psicología Educativa y Didáctica de la Lengua, la Literatura y las Artes [Educational Psychology and Didactics of Language, Literature, and the Arts] (E042-08, University of Murcia).

### References

- Aleksic, D., Cerne, M., Dysvik, A., & Skerlavaj, M. (2016). Preference for creativity, perceived clear outcome goals, work enjoyment, and creative performance. *European Journal of Work and Organizational Psychology, 25*(3), 363-383. <https://doi.org/10.1080/1359432X.2015.1077809>
- Argondizzo, C. (2012). *Creativity and innovation in language education*. Peter Lang. <https://doi.org/10.3726/978-3-0351-0431-8>
- Artola, T., Barraca, J., Mosteiro, P., Ancillo, I., Poveda, B., & Sánchez, N. (2012). *PIC-A. Prueba de Imaginación Creativa para Adultos*. TEA Ediciones. <https://bit.ly/3O3GEtn>
- Boden, M.A. (2004). *The creative mind: Myths and mechanisms*. Routledge. <https://doi.org/10.4324/9780203508527>
- Boden, M.A. (2018). *Artificial Intelligence: A very short introduction*. Oxford. <https://doi.org/10.1093/acrade/9780199602919.001.0001>
- Bonami, B., Piazzentini, L., & Dala-Possa, A. (2020). Education, big data and artificial intelligence: Mixed methods in digital platforms. [Educación, Big Data e Inteligencia Artificial: Metodologías mixtas en plataformas digitales]. *Comunicar, 65*, 43-52. <https://doi.org/10.3916/C65-2020-04>
- Breton (Ed.) (2021). *Una Europa adaptada a la era digital: la comisión propone nuevas normas y medidas para favorecer la excelencia y la confianza en la inteligencia artificial*. Comisión Europea. <https://bit.ly/3xVJvez>
- Csikszentmihalyi, M. (1998). *Creatividad: El flujo y la psicología del descubrimiento y la invención*. Paidós. <https://bit.ly/3pyqoqi>
- Donolo, D., & Elisondo, R. (2007). Creatividad para todos. Consideraciones sobre un grupo particular. *Anales de Psicología, 23*(1), 148-151. <https://doi.org/10.6018/analesps>
- Edelvives (Ed.) (2023). *Edelvives se convierte en la primera editorial en integrar ChatGPT en su plataforma educativa*. <http://bit.ly/3HN2rSu>
- Else, H. (2023). Abstracts written by ChatGPT fool scientists. *Nature, (7944)*, 423-423. <https://doi.org/10.1038/d41586-023-00056-7>
- Flores-Vivar, J.M., & García-Peñalvo, F.J. (2023). Reflexiones sobre la ética, potencialidades y retos de la Inteligencia Artificial en el marco de la Educación de Calidad (ODS4). [Reflections on the ethics, potential, and challenges of artificial intelligence in the framework of quality education (SDG4)]. *Comunicar, 74*, 37-47. <https://doi.org/10.3916/C74-2023-03>
- Fyfe, P. (2022). How to cheat on your final paper: Assigning AI for student writing. *AI & Society*. <https://doi.org/10.1007/s00146-022-01397-z>
- Gade, R. (2023). *Become a published author using ChatGPT*. PUBLISHER S21598.
- Goleman, D., Kaufman, P., & Ray, M. (2016). *El espíritu creativo*. Ediciones B. <https://bit.ly/44WZpoD>
- Guilford, J.P. (1950). Creativity. *American Psychologist, 5*(9), 444-454. <https://doi.org/10.1037/h0063487>

- Gutiérrez-Fandiño, A., Armengol-Estapé, J., Pàmies, M., Llop-Palao, J., Silveira-Ocampo, J., Carrino, C.P., González-Aguirre, A., Armentano-Oller, C., Rodríguez-Penagos, C., & Villegas, M. (2022). MarIA: Spanish Language Models. *Procesamiento del Lenguaje Natural*, 68, 39-60. <https://doi.org/10.26342/2022-68-3>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education. Promises and implications for teaching and learning*. Center for Curriculum Redesign. [https://doi.org/10.1007/978-3-319-60013-0\\_107-1](https://doi.org/10.1007/978-3-319-60013-0_107-1)
- Hora, S., Badura, K.L., Lemoine, G.J., & Grijalva, E. (2022). A meta-analytic examination of the gender difference in creative performance. *Journal of Applied Psychology*, 107(11), 1926-1950. <https://doi.org/10.1037/apl0000999>
- Hutson, M. (2022). Could AI help you to write your next paper? *Nature*, 611, 192-193. <https://doi.org/10.1038/d41586-022-03479-w>
- INTEF (Ed.) (2022). *IA y educación: Orientaciones para los responsables de la elaboración de políticas*. Ministerio de Educación y Formación Profesional. <https://bit.ly/3meypiA>
- Lee, A. (2023). *What are large language models used for?* NVIDIA. <http://bit.ly/3KHkFH3>
- Miller, A.I. (2019). *The artist in the machine. The world of AI-powered creativity*. The MIT Press. <https://doi.org/10.7551/mitpress/11585.001.0001>
- Montero, L. (2023). *¿Somos un cerebro flotando en una cubeta?* [Fallo de sistema, Radio 3]. <https://bit.ly/3mlagf0>
- Munari, B. (2018). *Fantasia. Invención, creatividad e imaginación en las comunicaciones visuales*. Gustavo Gili. <https://bit.ly/3LVYf3P>
- OCDE (Ed.) (2019). *PISA 2021 Creative thinking framework (third draft)*. <https://bit.ly/3lXBdt6>
- OCDE (Ed.) (2021). *OECD Digital Education Outlook 2021: Pushing the frontiers with artificial intelligence, blockchain and robots*. <https://doi.org/10.1787/589b283f-en>
- Oconnor, S., & Chatgpt (2023). Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse? *Nurse Education in Practice*, 66, 103537-103537. <https://doi.org/10.1016/j.nepr.2022.103537>
- OpenAI (Ed.) (2023). *Classifier*. <http://bit.ly/3KHD4nd>
- Peirón, F. (2023). Nueva York prohíbe el ChatGPT en sus escuelas. *La Vanguardia*. <http://bit.ly/41wFCe0>
- Real Academia Española -RAE- (s.f.). *Proyecto Lengua Española e Inteligencia Artificial (LEIA)*. <https://bit.ly/3SFlj92>
- Renbarger, M. (2023). Esta emprendedora ha usado ChatGPT para escribir un libro infantil en 2 horas: quiere demostrar que la IA generativa puede ser una poderosa herramienta de aprendizaje. *Business Insider*. <http://bit.ly/3J0Jcqv>
- Rodrigo-Martín, I., Rodrigo-Martín, L., & Pérez-García, A. (2022). La creatividad como herramienta para comprender la educación: El papel de la creatividad como catalizador de la transformación de la educación. *Visual Culture Review*, 9(3), 1-12. <https://doi.org/10.37467/revvisual.v9.3533>
- Sadín, E. (2018). *La inteligencia artificial o el desafío del siglo*. Caja Negra Editora. <https://bit.ly/3MiBQzg>
- Sánchez, A. (2023). El chat GPT obliga a repensar los exámenes y las prácticas en las aulas murcianas. *La Opinión*. <http://bit.ly/3ZnoBkv>
- Selwyn, N., Rivera-Vargas, P., Passeron, E., & Miño-Puigercós, R. (2022). ¿Por qué no todo es (ni debe ser) digital? Interrogantes para pensar sobre digitalización, datificación e inteligencia artificial en educación. In P. Rivera-Vargas, R. Miño-Puigercós, & E. Passeron (Eds.), *Educación con sentido transformador en la universidad* (pp. 137-148). Octaedro. <https://doi.org/10.31235/osf.io/vx4zr>
- Sternberg, R.J. (1999). *Handbook of creativity*. Cambridge University Press. <https://bit.ly/3nUDtd2>
- Sun, Z., Anbarasan, M., & Kumar, D.P. (2021). Design of online intelligent English teaching platform based on artificial intelligence techniques. *Computational Intelligence*, 37(3), 1166-1180. <https://doi.org/10.1111/coin.12351>
- Turing, A.M. (1950). Computing Machinery and Intelligence. *Mind*, (pp. 433-460). <https://doi.org/10.1093/mind/LIX.236.433>
- UNESCO (Ed.) (2019). *International conference on Artificial intelligence and Education, Planning education in the AI Era: Lead the leap: Final report*. <https://bit.ly/3kEhqpb>
- UNESCO (Ed.) (2021). *International Forum on AI and the Futures of Education, developing competencies for the AI Era, 7-8 December 2020: Synthesis report*. UNESCO. <http://bit.ly/3VKXAac>
- UNESCO (Ed.) (2022a). *Recomendación sobre la ética de la inteligencia artificial*. <https://bit.ly/3KBvPx8>
- UNESCO (Ed.) (2022b). *K-12 AI curricula: A mapping of government-endorsed AI curricula*. <https://bit.ly/3B6f6xi>
- UNESCO (Ed.) (s.f.). *Teaching AI for K-12*. <https://bit.ly/3KHl6A7>
- Universidad de Murcia (Ed.) (2022). *Código de Buenas Prácticas en Investigación de la Universidad de Murcia*. <http://bit.ly/3M1tZb>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30. <http://bit.ly/3kx6F89>
- Vázquez, D. (2023). Lo quieran los profesores o no, ChatGPT ya está aquí: Por qué muchos creen que hay que adoptar la IA en lugar de rechazarla. *Business Insider*. <http://bit.ly/3lxlsl1>
- Ward, M. (2020). Victorian fictions of computational creativity. In *AI Narratives: A history of imaginative thinking about intelligent machines* (pp. 144-164). Oxford University Press. <https://doi.org/10.1093/oso/9780198846666.001.0001>
- Yanhua, Z. (2020). *The application of artificial intelligence in foreign language teaching*. 2020 International Conference on Artificial Intelligence and Education (ICAIE). <https://doi.org/10.1109/ICAIE50891.2020.00017>