# Integrating Generative Language Models in Lesson Planning: A Case Study

Daša Berglez, Maja Kerneža
University of Maribor, Faculty of Education, Maribor, Slovenia
dasa.berglez@student.um.si, maja.kerneza1@um.si

The study explores artificial intelligence (AI) as a supportive technology in the lesson planning process of teachers, focusing on generative language models like ChatGPT. It evaluates the impact of AI on teaching strategies by comparing control and experimental groups of pre-service teachers interacting with the ChatGPT model. The aim was to create a lesson in metaphorical creativity and to understand the adaptability and effectiveness in lesson creation. Findings reveal no significant differences in pedagogical outcomes between groups, indicating that while AI offers innovative approaches, it doesn't fundamentally alter teaching effectiveness. The study It points out the potential and challenges of AI in education, advocating for further exploration to maximize its benefits and address ethical implications. This study contributes to the ongoing discourse on leveraging AI to enrich educational practices and emphasizes the critical role of educators in adapting AI tools for teaching.

**DOI**
https://doi.org/10.18690/um.pef.2.2024.11

**ISBN**
978-961-286-899-4

**Keywords:**
artificial intelligence, curriculum, educational technology, metaphorical creativity, teacher support

University of Maribor Press

**DOI**
https://doi.org/10.18690/um.pef.2.2024.11

**ISBN**
978-961-286-899-4

**Ključne besede:**
izobraževalna tehnologija,
metaforična ustvarjalnost,
učni načrt,
podpora učitelju,
umetna inteligenca

Univerzitetna založba
Univerze v Mariboru

# VKLJUČEVANJE GENERATIVNIH JEZIKOVNIH MODELOV V NAČRTOVANJE POUKA: ŠTUDIJA PRIMERA

DAŠA BERGLEZ, MAJA KERNEŽA
Univerza v Mariboru, Pedagoška fakulteta, Maribor, Slovenija
dasa.berglez@student.um.si, maja.kerneza1@um.si

Študija proučuje uporabo umetne inteligence kot podporne tehnologije v procesu načrtovanja pouka s strani učiteljev, pri čemer se osredotoča na generativne jezikovne modele, kot je ChatGPT. Namen študije je oceniti vpliv umetne inteligence na učne strategije z analizo in s primerjavo med kontrolnimi in eksperimentalnimi skupinami bodočih učiteljev. Eksperimentalna skupina je za načrtovanje učne ure o metaforični ustvarjalnosti uporabila ChatGPT z namenom, da bi ocenili njegovo prilagodljivost in učinkovitost. Rezultati ne kažejo pomembnih razlik v izidih med skupinami, kar nakazuje, da umetna inteligenca, čeprav prinaša inovativne pristope, ne spreminja osnovne učinkovitosti poučevanja. Študija izpostavlja pomen izobraževanja in vključevanja učiteljev pri uporabi umetne inteligence v izobraževalnem procesu. Osvetljuje potenciale in izzive umetne inteligence v izobraževanju ter spodbuja nadaljnje raziskave za izkoriščanje njenih prednosti. Rezultati študije prispevajo k razpravi o uporabi umetne inteligence za izboljšanje izobraževalnih praks in poudarjajo ključno vlogo učiteljev pri prilagajanju orodij umetne inteligence.

## 1 Introduction

Since late 2022, artificial intelligence (henceforth AI) has significantly impacted education, highlighted by the emergence of generative language model (henceforth GLM) ChatGPT (e.g. Baidoo-Anu & Owusu Ansah, 2023; Bonner et al., 2023; Kohnke et al., 2023). These developments, building upon earlier AI applications from basic computers to humanoid robots and interactive chatbots, have transformed both administrative and pedagogical practices, making education more personalized and interactive (Chen et al., 2020; Chen et al., 2021; Youfei et al., 2020). ChatGPT, known for its conversational abilities, has been particularly noted for its role in facilitating nuanced human-machine communication and demonstrating the potential of AI to support a diverse educational needs (OpenAI, 2022).

The integration of AI in education necessitates a comprehensive understanding of its role not just as a technological tool but as part of a broader societal construct influencing and reshaping educational methodologies (An & Oliver, 2021; Dron, 2022). Factors such as educators' background, institutional culture, and attitudes towards technology significantly influence the adoption and effectiveness of AI in teaching (Raygan & Moradkhani, 2022; Serin & Bozdag, 2020). This underscores the need for professional development tailored to educators' specific contexts to effectively utilize AI technologies (Kerneža and Zemljak, 2023a). Furthermore, the impact of AI on teaching and learning extends to influencing student success, with technology-assisted instruction showing positive effects on achievement, teaching efficacy, and engagement (Chaunan, 2016; Kareem, 2022). However, it's imperative to acknowledge the diversity among learners, as certain groups may derive more benefit from technology-supported education than others (Bergdahl et al., 2020). The decisions to incorporate chatbots in teaching, for example, is largely motivated by their perceived ease of use and usefulness, highlighting their potential to enhance educational practices (Choccarro et al., 2023).

Yet, for AI technologies to be effectively integrated into educational practices, educators must develop and refine a specific skill set. This includes the ability to identify and address technological challenges, think abstractly, and deeply understand and critically analyze the capabilities and outputs of AI tools like ChatGPT. Importance of such competencies, along with the broader implications of AI in education, including ethical, privacy, and reliability concerns, has been

emphasized in recent research (Kerneža, 2023; Vintar, 2023)- This underscores the critical role of ongoing teacher education and highlights the influence of teachers' self-assessed technology-related skills on the incorporation of contemporary technologies into educational activities (Chuan & Kabilan, 2021; Sailer et al., 2021).

Research indicates that AI, particularly through GLMs, significantly impacts educational outcomes across all levels, enhancing skills such as reading, writing, and critical thinking (Kasneci et al., 2023). These technologies support differentiate learning, accommodate diverse learner needs, and facilitate adaptive and inclusive educational environments (Jauhiainen & Guerra, 2023). However, the effective use of AI in education not only demands ongoing technological refinement but also a critical examination of its pedagogical integration to ensure it complements and enhances educational practices without compromising the quality of education or ethical standards.

*Research Problem*

While AI, particularly GLMs, holds transformative potential in education, it also introduces challenges and unanswered questions. This paper focuses on GLMs' role within the broad AI context in enhancing linguistic communication, a traditionally dominated by human interaction. Kunst Gnamuš (1984, p. 93) underscored communication's human-centric nature, driven by motivation and goals. ChatGPT's emergence as a tool of near-human communication, highlights the existing gap in AI's role in language skills development. Unlike human communication, the responses of the ChatGPT model are based on pre-existing data and lack the ability to adapt social needs and individual experiences as described by Kunst Gnamuš (1984). This highlights the need for specific research on the role of AI in developing communicative abilities, considering its limitations and potential in an evolving educational landscape.

*Research Focus*

This study explores the potential GLMs, like ChatGPT, to enhance communicative capabilities in education, focusing on metaphorical creativity among primary school students. Despite Lo (2023) highlighting the ChatGPT model's limitations in reliability and cross-subject effectiveness across, its contribution to teaching and

learning, particularly in linguistic and literary creativity, is substantial. This study delves in how AI facilitates the interplay of literal and figurative language, critical for developing children's metaphorical understanding as emphasized by Kunst Gnamuš (1984). Metaphors play a pivotal role in education, aiding in abstract concepts comprehension and reflecting the synthetic nature of human experience, pivotal for scientific innovation and diverse reality interpretations. The study underscores the necessity of fostering metaphorical expression in students to enhance their linguistic creativity through targeted exercises.

*Research Aim and Research Questions*

The primary aim of this study is to explore the utilization of AI, with specific emphasis on the GLM ChatGPT in the development of teaching strategies and the preparatory processes of the pre-service teachers. This research is centered on the realm of metaphorical creativity, seeking to examine the influence of GLMs on aspects such as pedagogical creativity and didactic coherence within the context of lesson planning. The research is guided by the following questions:

- How does the use of AI, specifically GLM ChatGPT, influence the pedagogical creativity in lesson plans developed by the pre-service teachers for metaphorical creativity?
- To what extent do lesson plans created by the pre-service teachers, with and without the assistance of AI, adhere to the theoretical foundations of metaphorical creativity? How do the pre-service teachers integrate AI into their lesson planning process, and what are the perceived limitations, considerations, and challenges of using AI, such as the ChatGPT model, in educational settings?

## 2 Research Methodology

*General Background*

This case study is based on a quasi-experimental design, fundamentally adopting both qualitative and quantitative approaches. It addresses identified gaps by exploring the impact of AI on the development of strategies and lesson preparations in the context of metaphorical creativity. This involves analyzing the interaction of

the pre-service teachers with the general language model ChatGPT and providing insights into the dynamics and challenges of integrating AI into educational processes.

*Sample*

The participants included in this study were 58 fourth-year pre-service teachers and the students from the Elementary Education program at one of Slovenian universities, specifically within the 'Didactics of Slovenian Language and Literature' course during the academic year of 2023/2024. As part of a study program, they are divided into four groups by their surnames. For the study, two of these groups were assigned as the control group and the other two as the experimental groups. These were further divided into 10 control (35 participants) and 8 experimental subgroups (23 participants) or lesson plan development. For analysis, these subgroups were treated as unified control and experimental group. All participant, having similar pedagogical backgrounds and AI exposure from the curriculum, were assumed to have a balanced distribution of prior teaching experience and AI familiarity. This grouping strategy, coupled with the exclusive use of ChatGPT 3.5 across the study, aimed to minimize confounding variables and attribute outcome differences directly to the use of AI in education.

*Instruments and Procedures*

In the Didactics of Slovenian Language and Literature course, pre-service teachers learn to develop primary school students' communicative competencies, including metaphorical creativity. They study theoretical foundations and practical approaches, culminating in creating lesson plans. The first part of the course was uniform for all, focusing on the theory and didactics of metaphorical creativity. In the second part, the participants were split into control and experimental groups. The control group created their lesson plans for metaphorical creativity independently, while the experimental group developed theirs with the assistance of AI, specifically the ChatGPT model. This interactive approach allowed the experimental group to explore and integrate innovative didactic methods and AI-supported technological solutions into their lesson plans. Detailed instructions regarding interactions with the GLM were not provided; instead, the decision was left to the participants' discretionary judgment, ensuring they could freely explore

and utilize ChatGPT in a manner that best suited their individual pedagogical goals and the specific requirements of their lesson plans. This autonomy in the use of ChatGPT was instrumental in fostering a creative and exploratory learning environment, where pre-service teachers could experiment with and reflect upon the integration of AI technologies into educational settings. This approach also mimics real-world scenarios where educators might use AI tools without predetermined guidelines, thus reflecting a more organic interactions between the pre-service teachers and the technology. The designed lesson plans were compared from two aspects, in addition to exploring the peculiarities that emerged in working with general language model ChatGPT. In the second part, where the experimental groups were instructed to develop their lesson plans with the assistance of ChatGPT, detailed documentation of interactions was maintained. This included specific prompts given to ChatGPT, the model's responses, and how these responses influenced the lesson planning process. Each interaction was meticulously recorded to capture the nuances of how AI could contribute to educational content development. However, it is important to note that due to the extensive nature, specificity, and complexity of these interactions between teachers and ChatGPT, they are not addressed in the present article. The intricacy of these engagement with ChatGPT necessitates a separate a separate, in-depth study to explore the pedagogical implications and the potential transformative impact of AI in lesson planning.

The first aspect focused on various elements of pedagogical effectiveness, for which only the extreme points of the Likert scales are shown, with 1 indicating the minimal and 5 the maximal level:

- *Pedagogical creativity* – the inclusion of fresh and innovative approaches to learning metaphors that could stimulate greater student engagement.
- *Didactic coherence* – logical arrangement of activities, progression from simple to more complex concepts.
- *Interactivity* – the level of interactivity and encouragement of student collaboration.
- *Technology* – the use of technology and other digital resources.
- *Adaptability* – adaptation to different learning styles and abilities of students, flexibility of activities for differentiation.

- *In-depth understanding* – enabling a deeper understanding of metaphors, not just recognition and use.
- *Reflection* – the design of concluding activities to reinforce knowledge and encourage reflection on the learning material.

The second aspect evaluated adherence to the theoretical foundations of metaphorical creativity, as defined by Olga Kunst Gnamuš (1984), comparing how the experimental and control groups applied these principles, with only the extreme points of the Likert scales shown, where 1 is minimal and 5 is maximal adherence:

- *Understanding of theoretical foundations* – the extent to which participants demonstrated an understanding of the theoretical concepts of metaphorical creativity.
- *Application of theoretical knowledge in practice* – the use of theoretical knowledge in designing lesson plans.
- *Integration of theory into the teaching process* – the integration of theoretical knowledge into the planning and execution of the lesson.

The third aspect of the study examined the complexities of using AI in education, focusing on integration, limitations, and ethical considerations, as well as its adaptability and future trends. Insights were gained from the interactions of the pre-service teachers with the ChatGPT model, using a qualitative approach to analyze how they incorporate Ai into lesson planning and address its challenges. The analysis utilized transcriptions of the conversations between participants and the ChatGPT model which were then systematically analyzed. The analysis of these interactions provided a deeper understanding of the role and potential of AI in educational settings.

*Data Analysis*

The data analysis aimed to understand, interpret, and connect the data, involving both qualitative and quantitative methods. Annotations were added, data were categorized and prepared for coding. Manual coding was conducted, utilizing open coding to identify fundamental themes and axial coding to connect these themes. The coding process was repeatedly reviewed and verified for consistency by two

independent researchers. IBM SPSS Statistics 27 was used for statistical analysis, including a Chi-Square test to test hypotheses regarding differences in pedagogical effectiveness and theoretical foundations among different student groups. The study also employed thematic analysis to explore AI nuances in education, considering methodological limitations and potential biases.

The qualitative analysis process commenced with the systematic examination of transcripts from the interactions between participants and ChatGPT. Initially, manual coding was applied using an open coding technique to discern primary themes. Subsequently, axial coding linked these themes to form a coherent understanding of the data. This iterative coding process was rigorously reviewed for consistency and reliability by two independent researchers, ensuring methodological rigor. The coding scheme was devised to categorize the types of pedagogical strategies derived from interactions with ChatGPT, focusing on creativity, adaptability, and technological integration. Categories were aligned with the research objectives to elucidate how ChatGPT could augment traditional lesson planning approaches.

This study was conducted adhering to the highest ethical standards, ensuring all participant interactions with ChatGPT were anonymized and used solely for the purpose of this research. Participants were informed about the research objectives and consented to their data being used for analysis. It was conducted in accordance with the research standards and ethics of Institute of Contemporary Technology, Faculty of Natural Science and Mathematics, University of Maribor (FNM_ICT) and approved by the Ethical commission for studies involving humans (1_2022).

## 3 Research Results

*Pedagogical Effectiveness*

The results of pedagogical effectiveness demonstrated by participants in the control and experimental groups are shown in Table 1.

**Table 1: Comparison of Results Between the Control and Experimental groups based on Demonstrated Pedagogical Effectiveness**

| | Rating | Control Group | | Experimental Group | | Chi-Square Test |
|---|---|---|---|---|---|---|
| | | f | f % | f | f % | |
| **Pedagogical creativity** | 1 | 0 | 0.0 | 0 | 0.0 | $\chi^2$ (3) =2.292, p = .514 |
| | 2 | 1 | 10.0 | 1 | 12.5 | |
| | 3 | 5 | 50.0 | 2 | 25.0 | |
| | 4 | 3 | 30.0 | 2 | 25.0 | |
| | 5 | 1 | 10.0 | 3 | 37.5 | |
| **Didactic coherence** | 1 | 0 | 0.0 | 0 | 0.0 | $\chi^2$ (2) =3.420, p = .181 |
| | 2 | 0 | 0.0 | 0 | 0.0 | |
| | 3 | 3 | 30.0 | 2 | 25.0 | |
| | 4 | 7 | 70.0 | 3 | 37.5 | |
| | 5 | 0 | 0.0 | 3 | 37.5 | |
| **Interactivity** | 1 | 0 | 0.0 | 0 | 0.0 | $\chi^2$ (3) =2.340, p = .505 |
| | 2 | 0 | 0.0 | 1 | 12.5 | |
| | 3 | 4 | 40.0 | 2 | 25.0 | |
| | 4 | 4 | 40.0 | 2 | 25.0 | |
| | 5 | 2 | 20.0 | 3 | 37.5 | |
| **Technology** | 1 | 0 | 0.0 | 2 | 25.0 | $\chi^2$ (3) =2.292, p = .514 |
| | 2 | 6 | 60.0 | 3 | 37.5 | |
| | 3 | 3 | 30.0 | 3 | 37.5 | |
| | 4 | 0 | 0.0 | 0 | 0.0 | |
| | 5 | 1 | 10.0 | 0 | 0.0 | |
| **Adaptability** | 1 | 0 | 0.0 | 0 | 0.0 | $\chi^2$ (3) =4.621, p = .202 |
| | 2 | 0 | 0.0 | 1 | 12.5 | |
| | 3 | 5 | 50.0 | 3 | 37.5 | |
| | 4 | 5 | 50.0 | 2 | 25.0 | |
| | 5 | 0 | 0.0 | 2 | 25.0 | |
| **In-depth understanding** | 1 | 0 | 0.0 | 0 | 0.0 | $\chi^2$ (2) =2.957, p = .228 |
| | 2 | 0 | 0.0 | 0 | 0.0 | |
| | 3 | 4 | 40.0 | 3 | 37.5 | |
| | 4 | 6 | 60.0 | 3 | 37.5 | |
| | 5 | 0 | 0.0 | 2 | 25.0 | |
| **Reflection** | 1 | 0 | 0.0 | 0 | 0.0 | $\chi^2$ (3) =7.513, p = .057 |
| | 2 | 0 | 0.0 | 1 | 12.5 | |
| | 3 | 3 | 30.0 | 4 | 50.0 | |
| | 4 | 7 | 70.0 | 1 | 12.5 | |
| | 5 | 0 | 0.0 | 2 | 25.0 | |

The analysis of Table 1 reveals no statistically significant differences in pedagogical effectiveness between the experimental and control groupsm as indicated by Chi-Square tests across various pedagogical aspects. Notably, both groups showed similar levels of innovation in teaching metaphors, with no statistically significant advantage observed for either group $\chi^2$ (3) =2.292, p = .514), although the experimental group tended to score higher. In examining didactic coherence, which

assesses the logical arrangement of activities, the experimental group displayed marginally better outcomes. However, these differences did not achieve statistical significance ($\chi^2$ (2) =3.420, p = .181), suggesting that the use of ChatGPT does not markedly alter the structural coherence of lesson. Interactivity and encouragement of student collaboration also showed no statistically significant differences ($\chi^2$ (3) =2.340, p = .505), despite a slight tendency for higher ratings within the experimental group. This trend continued with the use of technology and digital resources, adaptability to various learning styles, and the design of activities for in-depth understanding and reflection. In each case, while the experimental group occasionally showed higher scores, none of these differences reached statistical significance, indicating that integrating ChatGPT into lesson planning does not significantly impact these aspects of pedagogical effectiveness (Technology: $\chi^2$ (3) =2.292, p = .514; Adaptability: $\chi^2$ (3) =4.621, p = .202; In-depth Understanding: $\chi^2$ (2) =2.957, p = .228; Reflection: ($\chi^2$ (3) =7.513, p = .057). Despite the absence of statistically significant differences, the experimental group's tendencies towards higher ratings in several pedagogical aspects warrant attention. This observation suggests that while the immediate statistical impact of using ChatGPT on pedagogical effectiveness may not be evident, there is potential for AI to subtly influence educational practices. It highlights the need for further exploration into how AI, when thoughtfully integrated, might contribute to enhancing various dimensions of teaching and learning, even in the absence of stark statistical evidence.

*Theoretical foundations of metaphorical creativity*

The extent to which participants in the control and experimental groups adhered to the theoretical foundations of metaphorical creativity is shown in Table 2.

The analysis between the control and experimental groups shows that utilizing AI, specifically general language models like ChatGPT, in lesson planning did not significantly influence the participants' grasp and application of the theoretical concepts underpinning metaphorical creativity. The statistical analysis, conducted via the Chi-Square test, revealed no statistically significant differences between the groups' understanding or application of these concepts (Understanding: $\chi^2$ (2) =2.831, p = .243; Application: $\chi^2$ (2) =.070, p = .966; Integration: $\chi^2$ (2) =2.831, p = .243). This suggests that both groups, irrespective of AI tool usage, exhibited

comparable proficiency in engaging with the core principles of metaphorical creativity.

**Table 2: Comparison of Results Between Control and Experimental Groups Based On Adherence to the Theoretical Foundations of Metaphorical Creativity.**

| | Rating | Control Group | | Experimental Group | | Chi-Square Test |
|---|---|---|---|---|---|---|
| | | f | f % | f | f % | |
| **Understanding of theoretical foundations** | 1 | 0 | 0.0 | 0 | 0.0 | $\chi^2$ (2) =2.831, p = .243 |
| | 2 | 0 | 0.0 | 2 | 25.0 | |
| | 3 | 7 | 70.0 | 4 | 50.0 | |
| | 4 | 3 | 30.0 | 2 | 25.0 | |
| | 5 | 0 | 0.0 | 0 | 0.0 | |
| **Application of theoretical knowledge in practice** | 1 | 0 | 0.0 | 0 | 0.0 | $\chi^2$ (2) =.070, p = .966 |
| | 2 | 1 | 10.0 | 1 | 12.5 | |
| | 3 | 6 | 60.0 | 5 | 62.5 | |
| | 4 | 3 | 30.0 | 2 | 25.0 | |
| | 5 | 0 | 0.0 | 0 | 0.0 | |
| **Integration of technology into the teaching process** | 1 | 0 | 0.0 | 0 | 0.0 | $\chi^2$ (2) =2.831, p = .243 |
| | 2 | 0 | 0.0 | 2 | 25.0 | |
| | 3 | 7 | 70.0 | 4 | 50.0 | |
| | 4 | 3 | 30.0 | 2 | 25.0 | |
| | 5 | 0 | 0.0 | 0 | 0.0 | |

This equivalence in performance indicates a pivotal insight: the integration of AI in this context does not markedly enhance or diminish educators' theoretical engagement with metaphorical creativity. It points to the current state of AI tools as adjunct aids that do not fundamentally alter the pedagogical process concerning theoretical understanding and its practical application. Such findings highlight the necessity for educators to perceive AI technologies as complementary resources that support, rather than supplant, traditional educational methodologies. Given this understanding, it becomes imperative for future research to probe deeper into how AI can be optimized to contribute meaningfully to the theoretical and practical aspects of education. This exploration should not only assess the technological capacities of AI but also explore innovative pedagogical strategies that leverage AI to foster enriched educational outcomes.

*Nuances of Working with AI in the Process of Educational Planning*

In the third aspect of the study, the nuances of working with AI, particularly the ChatGPT model, were explored through detailed analysis of the conversations between the pre-service teachers and the aforementioned GLM. The analysis

revealed critical engagement with this technology, highlighting both the potential and limitations of AI in educational settings. For instance, it was observed that the assumptions of the ChatGPT model about student ages and the appropriate grade level for metaphor lessons indicate a need for precise input to generate accurate content. The study also notes a tendency for the GLMs to extend the duration of the lessons and occasionally provide definitions or examples that were not entirely suitable for the intended age group. These observations underscore the importance of critical assessment and adaptation when integrating AI into lesson planning.

Furthermore, the discussions revealed a recurring theme of the GLM's lack of familiarity with specific educational systems and terminologies, necessitating further customization for local contexts. Occasional grammatical errors and the need for time constraints in lesson planning with AI were also highlighted. Despite these challenges, there were instances where the ChatGPT model provided innovative ideas and alternative activities, demonstrating its potential as a tool for enhancing creativity and adaptability in lesson planning. However, the reliance on AI for group work and discussion prompts was noted, along with its variable definitions of metaphors, indicating the need for educators to critically evaluate and possibly modify the AI-generated content.

A significant concern identified was the appropriateness of the metaphors suggested by the ChatGPT model. Some metaphors were identified as either directly translated from other languages or entirely fabricated, lacking cultural relevance or comprehensibility in the targeted language. This not only diminishes the educational value of the examples provided but also underlines the challenges of using AI for nuanced language tasks like metaphor teaching. Additionally, the complexity of definitions and examples provided by the ChatGPT model was often beyond the suitable level for the students, suggesting a gap in the GLM's ability to tailor content to varying cognitive and developmental stages.

Overall, the interactions with the ChatGPT model provided valuable insights into the dynamic nature of integrating AI in educational processes. While it offers innovative approaches and can support various aspects of lesson planning, its effective use requires awareness of its limitations and a critical approach to integrating its suggestions into teaching practice. This part of the study contributes to a broader understanding of how AI can be harnessed to enrich educational

strategies, while also stressing the critical role of educators in mediating and contextualizing AI-generated content. The need for careful selection and adaptation of AI-generated metaphors and definitions is particularly emphasized, ensuring that they are age-appropriate, culturally relevant, and pedagogically sound.

## 4 Discussion

The study of AI's integration, particularly via the ChatGPT model in educational settings, highlights a nuanced understanding the role of Ai in education. Despite the lack of significant differences in pedagogical effectiveness between AI-utilizing and non-utilizing groups, the study reveals AI's capacity to innovate pedagogical strategies without fundamentally altering core outcomes. The occasional higher creativity ratings in the experimental group, although not statistically significant, echo the importance of creativity in education as discussed by Oktradiksa et al. (2021) and Patston (2021). This aligns with Boden's (1998) view of creativity as inherent to human intelligence and challenges posed to AI, and Jennings' (2010) notion of 'creative autonomy' in AI systems.

Significantly, the findings contribute to the discourse on AI's practical application in education, suggesting that AI, like the ChatGPT model. Holds considerable potential for fostering innovative teaching strategies and personalized learning experiences, as seen in the non-significant yet educationally meaningful enhancements in pedagogical creativity. This potential underlines AI's role in offering novel perspectives and methods that traditional statistical methods may not fully capture, emphasizing the need for further exploration of AI's long-term educational impacts, including student learning outcomes and skill development (Vintar, 2023).

The study also underscores the indispensable role of educators in the AI integration process. Despite AI's adaptability in suggesting diverse activities, the effectiveness of these AI-generated suggestions heavily relies on educators' ability to integrate them into the curriculum thoughtfully. This highlights a critical need for precise and context-specific input to AI systems to ensure the relevance and appropriateness of AI-generated content, addressing potential misinterpretations or generalizations by AI regarding student ages and lesson suitability. Addressing AI's limitations in understanding specific educational contexts necessitates scalable frameworks and

collaborative efforts to enhance AI tools' adaptability, such as advanced contextual adaptation algorithms and platforms for educator-AI developer collaboration. These efforts are pivotal for leveraging AI's educational vale effectively and aligning its deployment with pedagogical objectives. Ethical considerations, including biases and privacy concerns associated with AI use in education, demand a thoughtful approach to ensure ethical AI integration, balancing technological advances with pedagogical ethics and social responsibility. The need for further research is evident, with future studies advised to broaden demographic and contextual variable analyses and explore AI's long-term effects on education.

In a future deeply intertwined with AI (Kordigel Aberšek & Aberšek, 2020), the imperative role of AI in reshaping education is undeniable (Grassini, 2023). This study delves into AI's current and potential impact in education, highlighting ChatGPT's capabilities. It reaffirms the educator's indispensable role in leveraging AI to enrich pedagogical practices and emphasizes the necessity for ongoing adaptation and development of teaching strategies that incorporate AI technologies effectively. AI's integration in education suggests a paradigm shift where teaching strategies, student outcomes, and critical thinking development are concerned. Educators are pivotal in this transformation, guiding the seamless incorporation of AI into the curriculum to augment traditional teaching methods. This requires a profound understanding of AI's functionalities and limitations, ensuring Ai's contributions are both meaningful and ethically sound. The dynamic between AI tools like ChatGPT and educational content creation showcases the potential for customized learning experiences, addressing the diverse needs of students and promoting deeper engagement and understanding.

Grounded in a case study approach, this study's quasi-experimental design with 58 participants exposes limitations in statistical power and generalizability, especially in identifying minor effects of AI integration in educational settings. The focus on ChatGPT, while insightful, restricts the examination to a narrow slice of the AI spectrum, thus curtailing the findings' wider applicability. The research context and potential biases, such as selection and confirmation biases, further condition the applicability of our results beyond the specific sample explored. These factors underline the necessity for expansive and diverse future research to comprehensively grasp the potentials and hurdles of AI in education, advocating for a broader exploration of AI tools and a commitment to methodological rigor.

The IMPACT model Kerneža and Zemljak (2023b) exemplifies a structured approach to adopting AI in teaching, highlighting the importance of strategic planning in AI integration to achieve pedagogical objectives. As AI evolves, it presents an expansive opportunity to enhance educational practices, contingent upon its judicious and critical application aligned with educational ethics and goals. Flogie and Aberšek (2021, p. 97) caution against uncritical adoption of AI, advocating for a balanced approach that considers ethical implications and the value AI brings to the educational process. Furthermore, it is critical to recognize not every AI application is inherently beneficial or ethically acceptable. The discernment of AI's utility and ethical deployment in education should engage a broad spectrum of expertise beyond the technological domain, incorporating insights from behavioral sciences to ensure AI's impact is both positive and equitable.

This discussion reiterates the significance of AI in modern education, advocating or a thoughtful approach that leverages AI's strengths while addressing its challenges. It calls for an educator-driven integration process, underpinned by frameworks like the IMPACT model, to navigate the complexities of AI in education. The future of education with AI is not only about embracing new technologies but also about fostering an inclusive, innovative, and ethically responsible learning environment.

## 5 Conclusions

The exploration of AI, particularly GLMs like ChatGPT, is crucial in reshaping education. The main conclusion of this study emphasizes the necessity of balanced integration of technology in education. It emphasizes that the true value of AI lies not in replacing the educator but in augmenting the educational experience with innovative strategies and resources.

This research provides a comprehensive analysis of the nuanced role of AI in education, offering insights into its potential limitations. It concludes that the integration of AI into educational planning is a complex, multifaceted process that requires careful consideration, critical assessment, and strategic implementation. The findings of this study contribute to a more complete understanding of how AI can be effectively harnessed to enrich educational strategies, ensuring that it serves as a tool that aligns with and enhances pedagogical goals.

As we stand at this milestone of the future of education, marked by the rapid advancement of AI, it is important that educators, policymakers, researchers and others recognize both the immense potential and the inherent challenges of integrating AI into the educational setting. The journey of integrating AI into education is ongoing, and this study provides one step towards understanding its trajectory. It is a call to action for continued exploration, critical evaluation, and thoughtful implementation of AI in educational settings. The future of education is not predetermined. It is shaped by our collective efforts to understand, adapt, and innovate. As we all continue to navigate the evolving landscape of AI in education, this study serves as a foundation, emphasizing the need for a nuanced approach to integrating technology in pedagogical practices. The path forward is one of collaboration, innovation, and ongoing inquiry to ensure that the integration of AI not only meets educational standards but also enriches the learning experience for all.

**Acknowledgments**

The authors want to thank the research program P5-0433; Digital Restructuring of Deficit Occupations for Society 5.0 (Industry 4.0). The research program is financed by the Slovenian Research Agency (ARRS).

**References**

An, T., & Oliver, M. (2021). What in the world is educational technology? Rethinking the field from the perspective of the philosophy of technology. *Learning, Media and Technology, 46*(1), 6–19. https://doi.org/10.1080/17439884.2020.1810066

Baioo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI, 7*(1), 52–62. https://doi.org/10.61969/jai.1337500

Bergdahl, N., Nouri, J., Fors, U., & Knutsson, O. (2020). Engagement, disengagement and performance when learning with technologies in upper secondary school. *Computer & Education, 149*, 103783. https://doi.org/10.1016/j.compedu.2019.103783

Boden, M. A. (1998). Creativity and artificial intelligence. *Artificial intelligence, 103*, 347–356.

Bonner, E., Lege, R., & Frazier, E. (2023). Large language mode-based artificial intelligence in the language classroom: Practical ideas for teaching. *Teaching English with Technology, 23*(1), 23–41. http://dx.doi.org/10.56297/BKAM1691/WIEO1749

Chang, C. H., & Kidman, G. (2023). The rise of generative artificial intelligence (AI) language models – challenges and opportunities for geographical and environmental education. *International Research in Geographical and Environmental Education, 32*(2), 85–89. https://doi.org/10.1080/10382046.2023.2194036

Chauhan, S. (2016). A meta-analysis oft he impact of technology on learning effectiveness of elementary students. Computers & Education, 105, 14–30. https://doi.org/10.1016/j.compedu.2016.11.005

Chen, L., Chen, P., and Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access, 8*, 75264–75278. https://doi.org/10.1109/access.2020.2988510

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational technology and Society, 25*(1), 28–47. https://www.jstor.org/stable/48647028

Chocaro, R., Cortinas, M., & Marcos-Matas, G. (2023). Teachers' attitudes towards chatbots in education: A technology acceptance model approach considering the effect of social language, bot proactiveness, and users' characteristics. *Educational Studies, 49*(2), 295–313. https://doi.org/10.1080/03055698.2020.1850426

Chuah, K. M., & Kabilan, M. K. (2021). Teachers' views on the use of chatbots to support English language teaching in a mobile environment. International *Journal of Emerging Technologies in Learning, 16*(20), 223–237. http://dx.doi.org/10.3991/ijet.v16i20.24917

Dron, J. (2022). Educational technology: What it is and how it works. *AI & Society, 37*. 155–166. https://doi.org/10.1007/s00146-021-01195-z

Flogie, A., & Aberšek, B. (2021). Artificial intelligence in education. In O. Lutsenko, & G. Lutsenko (Eds.), *Active Learning. Theory and Practice* (pp. 97–118). IntechOpen. http://dx.doi.org/10.5772/intechopen.96498

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences, 13*(7), 692. https://doi.org/10.3390/educsci13070692

Jauhiainen, J. S., & Guerra, A. G. (2023). Generative AI and ChatGPT in school children's education: Evidence from a school lesson. *Sustainability, 15*(18), 14025. https://doi.org/10.3390/su151814025

Jennings, K. E. (2010). Developing creativity: Artificial barriers in artificial intelligence. *Minds and Machines, 20*, 489–501. https://doi.org/10.1007/s11023-010-9206-y

Kareem, J., Thomas, R. S., & Nandini, V. S. (2022). A conceptual model of teaching efficacy and beliefs, teaching outcome expectancy, student technology use, student engagement, and 21st-century learning attitudes: A STEM education study. Interdisciplinary Journal of Environmental and Science Education, 18(4), e2282. https://doi.org/10.21601/ijese/12025

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gassser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., …, Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103,* 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kerneža, M. (2023). Fundamental and basic cognitive skills required for teachers to effectively use Chatbots in education. In V. Lamanauskas(Ed.), *Science and technology education: New developments and Innovations. Proceedings of the 5th International Baltic Symposium on Science and Technology Education (BalticSTE2023)* (pp. 99-110). Scientia Socialis Press. https://doi.org/10.33225/BalticSTE/2023.99

Kerneža, M., & Zemljak, D. (2023a). Science teachers' approach to contemporary assessment with a reading literacy emphasis. *Journal of Baltic Science Education, 22*(5), 851–864. http://dx.doi.org/10.33225/jbse/23.22.851

Kerneža, M., & Zemljak, D. (2023b). IMPACT model: A comprehensive framework for planning the use of chatbots in didactic unit. In L. Gómez Chova, A. López Martínez, J. Lees (Eds.), *ICERI2023 Proceedings. 16th International Conference on Education, Research and Innovation, November 13th–15th, 2023, Seville, Spain*, pp. 1522–1527. IATED Academy. https://doi.org/10.21125/iceri.2023.0477

Kohnke, L., Moorhouse, B. M., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Kordigel Aberšek, M., & Aberšek, B. (2020). *Society 5.0 and Literacy 4.0 for the 21st century*. Nova Science Publishers. https://doi.org/10.52305/ZJBJ8012

Kunst-Gnamuš, O. (1984). *Govorno dejanje – družbeno dejanje: komunikacijski model jezikovne vzgoje [Speech act – social act: A communicative model of language education]*. Pedagoški inštitut.

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4): 410. https://doi.org/10.3390/educsci13040410

Oktradiksa, A., Bhakti, C. P., Kurniawan, S. J., Rahman, F. A., & Ani. (2021). Utilization artificial intelligence to improve creativity skills in society 5.0. *Journal of Physics: Conference Series, 1760*(1), 012032.https://doi.org/10.1088/1742-6596/1760/1/012032

OpenAI. (2022, November 30). *Introducing ChatGPT*. https://openai.com/blog/chatgpt

Patston, T. J., Kaufman, J. C., Cropley, A. J., & Marrone, R. (2021). What is creativity in education? A qualitative study of international curricula. *Journal of Advanced Academics, 32*(2), 207–230. https://doi.org/10.1177/1932202X20978356

Raygan, A., & Moradkhani, S. (2022). Factors influencing technology integration in an EFL context: Investigating EFL teachers' attitudes, TPACK level, and educational climate. *Computer Assisted Language Learning, 35*(8), 1789–1810. https://doi.org/10.1080/09588221.2020.1839106

Sailer, M., Stadler, M., Schultz-Pernice, F., Franke, U., Schoffmann, C., Paniotova, V., Husagic, L., & Fischer, F. (2021). Technology-related teaching skills and attitudes: Validation of a scenario-based self-assessment instrument for teachers. *Computers in Human Behavior, 115*(2), 106625. http://dx.doi.org/10.1016/j.chb.2020.106625

Serin, H., Bozdag, F. (2020). Relationship between teachers' attitudes towards technology use in education and autonomy behaviors. Turkish Online Journal of Educational Technology, 19(3), 60–69.

Vintar, Š. (2023). Jezik in umetna inteligenca: kam nas vodijo veliki jezikovni modeli [Language and artificial intelligence: Where big language models are taking us]. In J. Vogel (Ed.), *59. seminar slovenskega jezika, literature in kulture: Slovenski jezik, literatura, kultura in digitalni svet(ovi)*, pp. 21–28. https://centerslo.si/wp-content/uploads/2023/06/59.-SSJLK_Zbornik.pdf#page=21

Yufei, L., Saleh, S., Jahui, H., & Syed Abdullah, S. M. (2020). Review of the application of artificial intelligence in education. *International Journal of Innovation, Creativity and Change, 12*(8), 548–562.