Συνέδρια της Ελληνικής Επιστημονικής Ένωσης Τεχνολογιών Πληροφορίας & Επικοινωνιών στην Εκπαίδευση

14ο Συνέδριο ΕΤΠΕ «ΤΠΕ στην Εκπαίδευση»

14º Πανελλήνιο Συνέδριο με Διεθνή Συμμετοχή
"Οι ΤΠΕ στην Εκπαίδευση"

Τόμος Πρακτικών

Επιμέλεια

Σοφός, Α., Φωκίδης, Ε., Κώστας, Α., Φεσάκης, Γ., & Δημητρακοπούλου, Α.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ ΣΧΟΛΗ ΑΝΘΡΩΠΙΣΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ ΠΑΙΔΑΓΩΓΙΚΟ ΤΜΗΜΑ ΔΗΜΟΤΙΚΗΣ ΕΚΠΑΙΔΕΥΣΗΣ — ΕΤΠΕ ΕΛΛΗΝΙΚΗ ΕΠΙΣΤΗΜΟΝΙΚΗ ΕΝΩΣΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΑΣ & ΕΠΙΚΟΙΝΩΝΙΩΝ ΣΤΗΝ ΕΚΠΑΙΔΕΥΣΗ — ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ ΣΧΟΛΗ ΑΝΘΡΩΠΙΣΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ ΤΜΗΜΑ ΕΠΙΣΤΗΜΩΝ ΠΡΟΣΧΟΛΙΚΗΣ ΑΓΩΓΗΣ ΚΑΙ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ

ChatGPT as Pedagogical Agent in Higher Education: Insights into Reflective, Creative, and Critical Thinking

Sofia Hadjileontiadou, Ioannis Kapsidis, Eirini Tsetsiou

ChatGPT as Pedagogical Agent in Higher Education: Insights into Reflective, Creative, and Critical Thinking

Sofia Hadjileontiadou, Ioannis Kapsidis, Eirini Tsetsiou
schatzil@eled.duth.gr, dkapsidi@eled.duth.gr, etsetsio@eled.duth.gr
Department of Primary Education, Democritus University of Thrace

Abstract

In this study ChatGPT is embedded as pedagogical agent in the learning designs of two Case studies in Higher Education, involving 20 and 12 pre-service teachers respectively, to scaffold their creative and critical thinking. Thematic analysis of texts that were produced through reflective practices showed that students do not afford adequate competencies to distinguishing human from ChatGPT's creativity and to use critical thinking to evaluate its output for further use. Findings are not generalizable but raise awareness on necessary competencies when interacting with pedagogical agents.

Keywords: ChatGPT, creative thinking, critical thinking, pedagogical agent, the six hats of de Bono

Introduction

With the advancement of artificial intelligence and Large Language Models (LLM), chatbots have evolved into conversational agents (Aslan et al., 2025), capable of significantly mimicking human thought and speech, and engaging in open-ended dialogue (Sharma et al., 2025) on a wide range of topics in a natural and dynamic manner.

Naturally, the field of education could not remain unaffected by these developments. In educational contexts, conversational agents acquired the designation 'pedagogical' and are found a wide range of applications. They may assume the role of the teacher by providing instructional material, answering questions, generating content (Hwang & Utami, 2024), correcting exercises, and offering feedback and support. Alternatively, they may take on the role of the student in scenarios where teacher training in specific skills is required (Love et al., 2025). Moreover, LLM-powered conversational agents can, when provided with a relevant body of information, adopt fictional roles—such as that of a patient (Chan et al., 2025) for the training of medical professionals—or even take on multiple simultaneous roles (multi-agent), such as simulating a classroom of students, enabling user interaction with these simulated entities. In Rime's study (Rime, 2025), sixteen expert podcast profiles were created and queried on specific topics to gather data for defining product specifications. The study concludes that this tool can be effectively used to prepare for real interviews and that, with the appropriate input data, a coherent and convincing persona can be generated by the model, thereby circumventing issues related to privacy.

The use of pedagogical agents requires both critical and creative reflection. Critically, while such agents can increase engagement and motivation (Veletsianos & Russell, 2014), their impact on learning outcomes varies. Without thoughtful integration, they may lead to shallow learning or reduce student autonomy. Educators must examine not just whether agents work, but how and for whom they are effective. Creatively, pedagogical agents can support divergent thinking by prompting inquiry and problem-solving (Graesser et al., 2018). However, this depends on how interactions are designed. If agents offer rigid, scripted responses, they may limit rather than encourage creativity. To fully benefit from pedagogical agents, their implementation must be accompanied by ongoing critical analysis and intentional design to promote both deep learning and imaginative engagement.

In this work we embed ChatGPT in educational tasks, as pedagogical agent, to gain insights into creative and critical thinking of pre-service teachers while they use it. Although ChatGPT is not dedicated to education, we believe that the findings may contribute to some kind of sensitization of the interested in the research effort in pedagogical agents, as we hypothesize that the missing part in this dialogical procedure is the students' AI competencies and not the tool per se.

In the next section background information is provided. Then two cases of ChatGPT use are presented. In the following section the findings are discussed, and the conclusions and limitations of the work are provided.

Background information

Creative thinking

In the field of education, creativity appears to find a supportive environment (Morais et al., 2017) and can be developed across multiple subject areas throughout students' school life. This is particularly relevant for Higher Education Institutions (HEIs), which are responsible for preparing students to enter the labor market. These institutions should provide opportunities for students to cultivate creative skills, as at least one-third of them will be required to demonstrate specific creative competencies (Morais et al., 2017). The development of creative skills is important both on an individual and a broader societal level. Individually, it enables people to address everyday challenges more efficiently and contributes to an improved quality of life. On a societal level, creativity is vital for social progress, as it is linked to the capacity to transform new and innovative ideas into products and services. In a global context, creativity also plays a crucial role in addressing major challenges facing humanity in areas such as education, health, and the environment.

In literature the definition of creativity varies. In this work we adopt the well-known definition of Margaret Boden, who defines creativity as "the ability to produce ideas or artifacts that are new, surprising, and valuable" (Boden, 2004). She identifies three types of creativity: *combinational creativity*, which involves the unusual combination of existing ideas; *exploratory creativity*, which generates new ideas within an existing conceptual framework; and *transformational creativity*, which creates new ideas through a change in the conceptual space itself. According to Franceschelli and Musolesi (2024), Boden's model uses *novelty*, *surprise*, and *value* as the core criteria for judging creative outcomes. *Novelty* refers to something new and not previously existing, which can be framed psychologically (P-creativity) or historically (H-creativity) depending on the scope, individual or in humanity level, respectively. *Surprise* relates to the unexpectedness of the result and varies by creativity type. Lastly, *value* refers to the usefulness, effectiveness, and social appeal of an idea or artifact, emphasizing both quality and acceptance within a broader cultural or social context.

In the advent of conversational agents that can generate free-form text on virtually any requested topic, the impression is created that this technology is inherently creative. A series of research work refers to measuring and comparing the creativity of AI driven agents to human creativity. In the study of Charness and Grieco (2024), tasks were distributed along a continuum between open-ended and closed-ended problems. Results indicate that humans significantly outperform AI in solving open-ended problems, whereas AI excels in closed-ended ones. In a systematic literature review and subsequent analysis related to Generative AI (GenAI) and creativity (Holzner et al., 2025), it was found that GenAI achieves creativity levels comparable to the average human. Moreover, human-GenAI collaboration produces

better outcomes than human performance alone. However, this collaboration is associated with a significant reduction in the diversity of ideas generated.

Critical thinking

Critical thinking has emerged as one of the most vital skills of the 21st century, as it enhances learners' ability to evaluate, interpret, and synthesize information with logical coherence and evidence-based reasoning (Partnership for 21st Century Skills, 2009). According to Facione (1990), critical thinking is a purposeful, reflective, and self-regulating process of evaluation, governed by logical consistency, justification, and openness to revision. It aims at making well-supported decisions and forming reasoned judgments, and it is cultivated within metacognitive frameworks that require self-observation and the regulation of one's own thinking (Paul & Elder, 2008). It represents an inner discipline of the mind, requiring the development of both metacognitive and reasoning skills.

The cultivation of critical thinking is inextricably linked to metacognitive skills, namely the individual's ability to observe and regulate their own cognitive processes (Flavell, 1979), as well as reasoning skills (Paul & Elder, 2008). Students with well-developed metacognitive and critical thinking skills are able to ask questions, challenge biases, identify logical fallacies, evaluate reliable evidence, and make decisions based on data rather than arbitrary impressions (Halpern, 2014). Therefore, critical thinking is not confined to an academic pursuit but is embedded in the core mission of education as a process of developing responsible, thoughtful citizens.

The ACER Critical Thinking Framework (Heard, 2025) provides a structured approach for developing and assessing critical thinking. It is organized into three strands: Clarifying and Interrogating Ideas, Evaluating Arguments and Evidence, and Making Decisions or Drawing Conclusions. Each strand includes three aspects that guide the specific cognitive and metacognitive skills involved. The first strand focuses on understanding and exploring ideas, encouraging learners to interpret concepts, ask meaningful questions, and express their reasoning clearly. The second strand emphasizes the ability to assess the quality and relevance of arguments and evidence, fostering logical analysis and critical evaluation. The third strand involves forming reasoned conclusions and making decisions, promoting justified and reflective thinking.

The Six Thinking Hats of de Bono

One of the key challenges that emerge from the aforementioned points, concerns the structured support of thinking, to ensure that students develop conscious awareness of the cognitive strategies they employ and to activate alternative approaches when necessary. Within this context, Edward de Bono's (1985) "Six Thinking Hats" model serves as a powerful cognitive tool, providing a structured system for intellectual exploration. Through systematic practice of different modes of thinking, the model contributes to the development of a holistic and multidimensional approach to decision-making and problem-solving, reinforcing students' critical and creative thinking. Within this framework, Edward de Bono's (1985) "Six Thinking Hats" model is proposed as a systematic methodology that enhances critical and creative thinking. According to de Bono, the model is based on the categorization of distinct cognitive functions, represented symbolically by six colored hats:

White: Neutral and objective information, focused on facts and verified data.
Red: Emotional response, expressing intuition, empathy, and personal feelings.

Black: Critical thinking and risk analysis, identifying weaknesses and limitations of an idea or proposal.

Yellow: Positive thinking and optimism, emphasizing benefits and opportunities.

Green: Creative thinking and innovation, encouraging the exploration of alternative solutions and original ideas.

Blue: Metacognitive regulation and coordination, overseeing the thinking process and shaping strategic decisions.

The pedagogical use of this model aims to deliberately activate specific modes of thinking, preventing the confusion or simultaneous involvement of multiple—and often contradictory—mental operations. Moreover, the model can support the development of metacognitive strategies such as planning, monitoring, and evaluating thought processes, particularly through the function of the blue hat. Especially valuable is the red hat's ability to isolate and release emotion, allowing for the explicit acknowledgment of emotional or value-laden aspects without requiring justification, before thinking returns to more analytical pathways. Its implementation in discussion scenarios, case studies, or decision-making tasks empowers both individual and group cognitive engagement.

Cace studies

Two small-scale cases of exploratory character, namely Case A and Case B, involving the use of ChatGPT as pedagogical agent by volunteering pre-service teachers in a Greek HEI, provided qualitative data to support our aim to have insights into their reflective, creative, and critical thinking while using ChatGPT.

Case A

The context. Pre-service teachers in the HEI follow an optional, one semester course on the design and development of a serious game. Towards this endeavor, they are introduced to Scratch programming language, to serious games (theory, mechanics, dynamics, types of serious games and examples) and to the 2D-ME model on the way of working. According to this model, they are to realize two views of themselves, i.e., the 1st and the 3rd view, and adopt an internal dialogue while working on the game. The 1st view models the student when s/he is working on the game design and development, whereas the 3rd view is the student's creative self. This latter self is expected to release the power of exploration in the creativity space and innovate in game design and development. The 2D-ME model mirrors this internal collaboration on the triangle notion of the activity theory, where the two selves are to collaborate with the objective to design and develop a prototype of the serious game. In this effort they use tools that mediate this collaboration, whereas there are critical points that the division of labor takes place between the two selves. The students are to reflect on this activity while they design and develop their game and record this dialogue in their diary while they work. Provoking questions that are given to them help them to distinguish their two self-views (e.g., Q2.1 What contradictions can you identify between your different roles?, Q4.2 How do you organize the different interpersonal roles across the prototyping process?). The initial idea of the game is the starting point and through circles of optimization, a series of versions of the prototype are produced until it converges to the desired final edition of the game that is set by each student. The 2D-ME model proposes a pure reflecting way of working and makes explicit the two self-views of the game designer/developer. More on the 2D-ME model can be found in (Hadjileontiadou, et al., 2023). It is apparent that the creative 3rd

person's view of work can be detected and analyzed through the diary, combined with the prototype's editions.

Methodological approach. Within this framework, 20 students from the 3rd year of their studies in the HEI were asked to design and develop individually, as their final assignment, a serious game on a concept of their choice upon the relevant primary school's curriculum. Moreover, they were allowed to use ChatGPT as they wished. The task was open to allow for exploration of the creativity space from the 3rd person's view. In this case ChatGPT was used as a pedagogical agent that could serve as a tool to mediate the collaboration between the 1st and the 3rd person's views and possibly contribute to the creative work of the latter. The optional use of GhatGPT should be also reported in the reflection diary. A thematic analysis of the diary content was expected to provide insights into the role of ChatGPT to creativity thinking. Of the total 20 students, one third of them used ChatGPT, mainly concerning help to solve coding problems. From the diary analysis of these students based on Boden's theoretical perspective, it was revealed that they could not discriminate ChatGPT's creativity from human creativity due to misconceptions upon type and depth of creativity, cognitive processes employed and motivation and inspiration (Hadjileontiadou, 2025). Considering this fact, at the end of the semester we conducted an open-question survey to record the way that all 20 students perceive the role of ChatGPT as a pedagogical agent to their creativity.

Boden's (2004) approach to creativity was used as sensitizing theory for the thematic analysis of the students' answers. Here we present the results from the answers to the question: *How could ChatGPT reinforce the design and development of a serious game in Scratch from the 3rd person's view?*

Results. From the elaboration of the students' answers, the following misconceptions (M) were revealed as compared to Boden's work on creativity (Boden, 2004).

M1. *Creativity emerges out of nothing* [*ChatGPT can give me an idea out of a sudden that I could never think of myself*]. Boden considers that "creativity is grounded on thought-processes and mental structures" (p. 13), rather than appearing unpredictably. "Although 'sudden illumination' may be a faithful description of how creativity appears to the creator, it cannot be the whole story. Intuition cannot consist merely in flashes of insight. That way, magic lies." (p.28). Moreover, even the ChatGPT is trained upon data and the patterns that are output depend on the prompt. So, attributing this spontaneous behavior of ChatGPT reveals a lack of the way it works.

M2. *Negative Impact on personal Creativity*. Several students express worry that relying on ChatGPT could 'damage' their personal creativity. One states: [*I try not to use it as much as possible because I believe that if I use it continuously my creativity is slowly lessened*]. In this case there is a misconception that creativity is limited and somehow may end. This approach lacks the idea that creative thinking is a competency that can be cultivated. As Boden states "Computers can come up with new ideas and help people to do so. Both their failures and their successes help us think more clearly about our own creative powers" (p.10), thus human agency to creativity is never removed unless the person chooses not to be engaged.

M3. *There is a dichotomy between structured thinking and creativity*. Some students state that [...*creativity is not about coding or technical issues. It is purely fantasy*]. Such ideas reflect the notion that creativity cannot be found in systematic disciplines where logic is a profound way of thinking. Boden however defined conceptual spaces as "structured styles of thoughts" (p.4) and that "we must know what conceptual spaces are and what sorts of mental processes could explore and modify them" (p.6). Boden's model explicitly argues that algorithmic thinking and structured problem-solving can be deeply creative. Moreover, she states that "My account of human creativity will call on many computational ideas. ….. They can be thought of as a

particular class of psychological ideas. As we shall see, they help us to understand not only how creativity can happen, but also what creativity is." (p.15).

M4. *Concerning ChaGPT's role*. A student wrote [*ChatGPT gives me everything ready, I do not have to think anymore*] suggesting that the tool is the primary actor in the creative process. This contradicts Boden's view that AI and human creativity should work in partnership, with AI supporting the expansion or transformation of conceptual spaces.

Case B

The context. The second case refers to pre-service teacher students who attended the optional course 'Educational software' in the second year of their studies in the HEI. Towards the end of the semester, reflections on the course were challenged. To broaden the scene about educational software, the work of Manovich (2013) was given as supplementary material to read. The current generation of pre-service teachers (Gen Z, often labeled as "digital natives"), may not inherently possess sophisticated digital literacy or critical evaluation skills (Ng, 2012). Many engage in multitasking and rapid content consumption, often relying on algorithmically curated sources without questioning their validity (Kirschner & De Bruyckere, 2017). It was noticed that the specific pre-service teacher students did not choose to read the supplementary material and preferred the idea of its rapid elaboration. To avoid a surfaced approach of using ChatGPT and instead use it as pedagogical agent, 12 volunteers out of the 39 students that attended the course were asked to use ChatGPT in a structured activity as follows.

Methodological approach. The activity took place in a face-to-face mode to be supervised. The students were asked to perform individually three steps.

Step 1. The students, wearing the hats of de Bono, were asked to comment on the ChatGPT as pedagogical agent per se (without the use of ChatGPT).

Step 2. They were asked to use ChatGPT, in a role play, i.e., as Mr. Manovich, who, wearing the hats of de Bono, was asked to comment on the term 'educational software', exclusively from the perspective of his specific work (Manovich, 2013). The students were free to formulate the exact prompt to ChatGPT, provided that they followed these given instructions.

Step 3. The students, having read the output of ChatGPT from the previous step, wore the six hats of de Bono to commend on the term "educational software" (without the use of ChatGPT).

The texts that were produced either by ChatGPT or the students were analyzed. The ACER theoretical framework of critical thinking was used for the qualitative content analysis of this raw material in Step 3.

Results. Initially content analysis of the outputs from Step 1 was performed per de Bono's hat.

White: *Neutral and objective information, focused on facts and verified data.* From the elaboration of the answers of the students, three themes were identified, i.e., ChatGPT provides a vast amount of information quickly and efficiently, the information it gives is not always accurate or reliable. [*it gives me plenty of information, either correct or mistaken*] and critical thinking and human judgment are necessary when using it. The phrases are statements that evolved from the students' responses without any justification. The only effort to provide facts on the themes' context are two excerpts, yet quite surfaced [*it is AI tool that functions upon some human-made coding*] and [*it uses a vast amount of information but existing one*].

Red: *Emotional response, expressing intuition, empathy, and personal feelings*. The themes that evolved from the answers while wearing this hat were: Absence of genuine empathy in AI and mixed or neutral feelings towards ChatGPT (e.g., Positive feeling of being supported and

concerns and fear about rapid development of AI). The answers under this hat are subjective and no further justification is provided.

Black: *Critical thinking and risk analysis, identifying weaknesses and limitations of an idea or proposal.* The evolved themes were, Risk of Inaccurate or Misleading Information, Loss of Critical Thinking, Reduction of Creativity and Effort-depedence and mental laziness, Built-in Bias and Manipulation Potential [*it is a machine and it is reasonable to make mistakes*], [*if it is not used with discretion and critical thinking, the use of the human brain and creativity may be limited*], [*If someone bases their knowledge on it and trusts it without suspicion, they will stop thinking critically, as they are no longer engaging in thought*], [*It manipulates our thinking and doesn't allow us to think independently, therefore our mind is no longer cultivated or expanded*]. The students also of this semester carry the same misconceptions as the students in Cace A, about the constrained amount of creativity that every person possesses and thus it will be diminished upon the ChatGPT use, neglecting the potentiality of the human brain upon its cultivation.

Yellow: *Positive thinking and optimism, emphasizing benefits and opportunities*. The themes that evolved from this perspective were, Fast and Easy Access to Information (quick access to wide range of resources), Support for Learning and Academic Work (organization, clarification, and restructuring of information), inclusive and accessible technology [*it gives correct information and creates positive outputs*], [*It can enhance our creativity, give us new ideas, and help us create a good result-without, however, the absence of critical thinking*]. It seems that the students connect creativity with critical thinking apart from misconceptions.

Green*: Creative thinking and innovation, encouraging the exploration of alternative solutions and original ideas*. From the analysis of the answers in this perspective, three themes emerged, i.e., ChatGPT as source of creativity and inspiration [*the creativity of AI tools is very significant, as we are talking about programs that are configured and programmed to access a large volume of data. In this way, they create and "think" quite creatively-sometimes even more than the average person*], Limitations of ChatGPT creativity as compared to human creativity [*ChatGPT is quite creative, but in reality, it is a machine and therefore a human creation. So, humans are more creative*], [*ChatGPT is not creative. It is software that responds based on data. That is not creativity*] and thirdly, Need for responsible and balanced use [*its use should be within reasonable limits*], [*it is a tool for inspiration but not a replacement for humans*], [*It helps with the gathering and organizing of thoughts because the mind can become overloaded*].

Blue: *Metacognitive regulation and coordination, overseeing the thinking process and shaping strategic decisions*. The following themes evolved from the students' answers. Cognitive Structuring and Organization, Idea Refinement and Transformation, Accessibility and Assistance [*It helps organize data, such as better organizing notes, summarizing texts, and maintaining word limits in texts*], [*It helps with the gathering and organizing of thoughts because the mind can become overloaded*], [*Especially, it can assist individuals who have some language difficulties, such as dyslexia*], [*It uses logic and information it collects, so in my opinion, it is not logical on its own*], [*It organizes our ideas, enriches them with more, and makes our ideas more accessible*].

At the end of Step 1, the students proceeded to Step 2. The text (namely Text A) that the students received across the hats in Step 2 was a mixture of combination of educational software with excerpts from Manevich's work, thus in some points it included false information as far as Manevich's ideas are concerned. It must be noted that in the specific work of Manovich there is not any reference to educational software as the approach to software and media is more general and at some point, more abstract. The basic ideas of Manovich in this work are (a) the digitalization of data is a new format for media that puts a dichotomy between data and media function (b) digital data are stored as content in files accessible only to knowable (c) a media can efficiently access data only through interfaces as

explicated by software (d) software is a social construct and is provided by manufacturers (e) the new ways of media access, distribution, analysis, generation, and manipulation all come from software.

After the students read the output from ChatGPT they were asked to proceed to Step 3. Analysis of the text (namely Text B), that the students produced across the hats at this step, revealed that all of them adopted false positionings from text A concerning Manovich's ideas about the software.

[*I believe that Manovich would be more interested in learning where the information comes from (e.g., from which digital medium, etc.) to understand what additional information is needed to fully grasp the problem and the topic itself. I think the concern expressed is absolutely reasonable, as the immediacy and human contact in the educational process are indeed lost*]. Manovich stated that the affordances of the interface, and not the media per se, defines the manipulation of data. From the first sentence it is evident that the student has not realized the main argument of Manovich in his paper. Moreover, misconception continues in the second sentence where it is transferred to the educational space.

[*Manovich would emphasize that when the learning environment is transformed into an interface-package, pedagogical freedom is limited by the algorithm*]. Here also the student has paraphrased the main idea of Manovich as even the word freedom is not included in his text. Moreover, Manovich refers to the flexibility that software can offer in the elaboration of the same data as opposed to media affordances in earlier years. Furthermore, the main issue that Manovich is arguing about is the relation between media and software. So, at the pedagogical plain, arguments, like the aforementioned by the students, could be made in relation to the media that may be used in the classroom and the role of software behind. Other transfers far away from Manovich's ideas belong to assumptions that may not be justified critically.

Discussion and conclusions

Yusuf et al. (2025) offered a categorization for the conversational agents, from the pedagogical point of view) on three tiers. This work used ChatGPT as AI conversational agent and the way it was used is mapped to this categorization as follows. In the first tier of pedagogical purposes, its use focused on cognitive purpose, i.e., to support reflective functions upon creativity and critical thinking concerning the use of the agent per se. In the second tier, the mode of study was hybrid (Case A) and face to face (Case B). As far as the third tier is concerned, the conversational agent's intent was at the Reflective skills & metacognitive subcategory where the agent promotes analytical, divergent, critical and/or creative thinking in learners to aid learning. Moreover, from the technical point of view, ChatGPT is a disembodied, chatbot (text based) agent, widely used and embedded in Cases A and B without any specific technological elaboration.

The aim of the presented work was to have insights to reflecting creative and critical thinking on the use of ChatGPT as a pedagogical agent. From the reflection perspective, the mapping makes explicit that the effort of this work was on the pedagogical part to "build the metacognitive functions of the students by promoting creative thinking, giving experiences or reducing the psychological natural responses" (Yusuf et al., 2025, p.826).

In particular, in both Cases A and B, the students were asked to reflect on their internal dialogue in (Case A), and their semester work on educational software (Case B). ChatGPT was embedded in these efforts to function as a mediating tool and as such, it provided the space for the realization of the creative and critical thinking of the participating students concerning its use as conversational agent.

From the creative thinking perspective, in Case A, ChatGPT was considered as a means to expand the students' combinational and exploratory (Boden, 2004) creativity. However, its use as an aid to programming moved the pedagogical purpose to another category of Yusuf et al. (2025) taxonomy, i.e., the instructional, as ChatGPT instructed them how to solve problems on coding in Scratch. The survey that followed made explicit a series of misconceptions about creativity as described above. Moreover, students of another year of studies, who participated in Case B, were also asked to comment on ChatCPT under the different perspectives of de Bono's six hats (Step 1). In particular, the green hat refers to creativity. From the elaboration of the results also on this case, it was found that the aforementioned misconceptions about creativity were present in both Case A and B student samples. It is evident that although students mention the creativity perspective in the interaction with ChatGPT, it seems that they lack the relevant competencies to implement the creativity endeavor even when they have the challenge and space to do so, as foreseen in the learning design of both Cases.

Finally, concerning critical thinking, the learning design of Case B foresaw the reinforcement of the students' critical thinking, through de Bono's framework and expansion of the space of thought by provoking their semester background in Educational software through Manovich's ideas. From the results it was made clear that although the students mention that the use of ChatGPT requires critical thinking, there is lack in the relevant competencies to do so. Based on the ACER's critical thinking framework (Heard, 2025), the students' texts were poor in all three strands of critical thinking, as they lacked logical justification. They accepted the output of ChatGPT from its role play, without assessing it upon Manovich's text which was available to them. Here we can hypothesize the role of Gen Z characteristics for fast consumption of information without checking its source and even denoting elements of the evolving human-AI form of trust (Pitts & Motamedi, 2025).

Concluding, it is noted that, due to the small scale of both cases, mainly in terms of the sample size of the participants, the findings of this research work are definitely not generalizable. Yet, this work contributes to the ongoing discussion about pedagogical agents for cognitive purposes, a category that, according to Yusuf et al. (2025), remains underrepresented in the literature. The work does not contribute to the technological aspect of conversational agents but to the way an agent, which is not dedicated to education, like ChatGPT, can be embedded in the learning design for cognitive purposes. We believe that the findings may contribute to some kind of sensitization of the interested in the research effort in pedagogical agents, as the missing part as we realized in this dialogical procedure is the students' AI competencies and not the tool per se. The question that remains open is no longer whether we can integrate pedagogical agents like ChatGPT into education, but rather how and through what kind of thinking we can leverage them pedagogically and responsibly.

References

Aslan, S., Alyuz, N., Li, B., Durham, L. M., Shi, M., Sharma, S., & Nachman, L. (2025). An early investigation of collaborative problem solving in conversational AI-mediated learning environments. *Computers and Education: Artificial Intelligence, 8,* 100393. https://doi.org/10.1016/j.caeai.2025.100393

Boden, M. A. (2004). *The creative mind: Myths and mechanisms.* Routledge.

Chan, K. Y., Yuen, T. H., & Co, M. (2025). Using ChatGPT for medical education: The technical perspective. *BMC Medical Education, 25*(1), 201. https://doi.org/10.1186/s12909-025-06785-9

Charness, G., & Grieco, D. (2024). Creativity and AI. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.4686415

De Bono, E. (1985). *Six Thinking Hats.* Little, Brown and Company.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction.* The Delphi Report.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*(10), 906-911. https://doi.org/10.1037/0003-066X.34.10.906

Franceschelli, G., Musolesi, M. (2024). On the creativity of large language models. *AI & Society*, *40*, 3785-3795. https://doi.org/10.1007/s00146-024-02127-3

Graesser, A. C., Hu, X., Nye, B. D., VanLehn, K., Kumar, R., Heffernan, C., & Baer, W. (2018). ElectronixTutor: an intelligent tutoring system with multiple learning resources for electronics. *International journal of STEM education, 5*, 1-21. https://doi.org/10.1186/s40594-018-0110-y

Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking* (5th ed.). Psychology Press.

Hadjileontiadou, S. (2025). Διακρίνοντας τη δημιουργικότητα ανθρώπου και μεγάλων γλωσσικών μοντέλων κατά την αλληλεπίδρασή τους [Discriminating human creativity from large language models during their interaction [in Greek]. *Proceedings of the 5th ELOER*. University of Peloponnese and ETPE.

Hadjileontiadou S, Dias SB, Hadjileontiadis L. (2023). A conceptual framework (2D-ME) for explaining self-first and self-third person Views of prototyping dynamics in serious games design: experimental case study *JMIR* Serious Games *11*, e41824 https://doi.org/10.2196/41824

Heard J., Scoular, C., Duckworth, D., Ramalingam, D., & Teo, I. (2025). *Critical thinking: Skill development framework* (2nd ed.). Australian Council for Educational Research. https://doi.org/10.37517/978-1-74286-752-6

Holzner, N., Maier, S., & Feuerriegel, S. (2025). *Generative AI and creativity: A systematic literature review and meta-analysis*. arXiv preprint. https://doi.org/10.48550/ARXIV.2505.17241

Hwang, W.-Y., & Utami, I. Q. (2024). Using GPT and authentic contextual recognition to generate math word problems with difficulty levels. *Education and Information Technologies, 29*(13), 1-29. https://doi.org/10.1007/s10639-024-12537-x

Kirschner, P. A., & De Bruyckere, P. (2017). The myths of the digital native and the multitasker. *Teaching and Teacher Education, 67*, 135-142. https://doi.org/10.1016/j.tate.2017.06.001

Love, R., Law, E., Cohen, P. R., & Kulić, D. (2025). Teaching a conversational agent using natural language: Effect on learning and engagement. *International Journal of Artificial Intelligence in Education, 2025,* 1-39. https://doi.org/10.1007/s40593-025-00461-1

Manovich, L. (2013). Media after software. *Journal of Visual Culture, 12*(1), 30-37. https://doi.org/10.1177/1470412912470237

Morais, M. D. F., Azevedo, I., Fleith, D. D. S., Alencar, E. M. L. S. D., Almeida, L. S., & Araújo, A. M. (2017). Teaching practices for creativity at university: A study in Portugal and Brazil. *Paidéia (Ribeirão Preto), 27*(67), 56-64. https://doi.org/10.1590/1982-43272767201707

Ng, W. (2012). Can we teach digital natives digital literacy? *Computers & Education, 59*(3), 1065-1078.

Paul, R., & Elder, L. (2008). *The miniature guide to critical thinking concepts and tools.* Foundation for Critical Thinking.

Partnership for 21st Century Skills. (2009). *A framework for twenty-first century learning*. Partnership for 21st Century Skills.

Pitts, G., & Motamedi, S. (2025). *Understanding human-ai trust in education*. arXiv preprint. https://doi.org/10.48550/arXiv.2506.0916

Rime, J. (2025). Interviewing ChatGPT-generated personas to inform design decisions. In H. Plácido Da Silva, & P. Cipresso (Eds.), *Computer-human interaction research and applications* (Vol. 2370, pp. 82-97). Springer Nature. https://doi.org/10.1007/978-3-031-82633-7_6

Sharma, S., Mittal, P., Kumar, M., & Bhardwaj, V. (2025). The role of large language models in personalized learning: A systematic review of educational impact. *Discover Sustainability, 6*(1), 243. https://doi.org/10.1007/s43621-025-01094-z

Veletsianos, G., & Russell, G. (2014). Pedagogical agents. In *Handbook of research on educational communications and technology* (pp. 759-769). Springer. https://doi.org/10.1007/978-1-4614-3185-5_61

Yusuf, H., Money, A. & Daylamani-Zad, D. (2025). Pedagogical AI conversational agents in higher education: a conceptual framework and survey of the state of the art. *Educational Technology Research and Development, 73*, 815-874. https://doi.org/10.1007/s11423-025-10447-4