AI in board Game-Based Learning

Andrea Tinterri*1*, Marilena di Padova *2*, Francesco Palladino*3*, Giordano Vignoli*4* and Anna Dipace*2*

1 IUL Telematic University, Via M. Buonarroti 10, 50122, Florence, Italy
2 University of Foggia, Via A.Gramsci 89/91, Foggia, Italy
3 University of Modena and Reggio Emilia, Via Università 4, 41121 Modena, Italy
4 IESS, Istituto Europeo Studi Superiori, P.zza Prampolini, 2/A, 42121 Reggio Emilia

Abstract

Despite receiving less attention in educational research compared to digital games, boardgames show great potential as a learning environment in many educational scenarios. They promote acquisition of disciplinary knowledge and key competences, generate a sense of physical "togetherness", can be employed in situation of social and economic disadvantage, and can be modified (or "modded"), for better alignment with disciplinary content. The use of games in European schools is very limited; teachers see the potential of games for learning, but their competence in the use of games for learning is superficial and limited to personal experience. High-performance AI systems such as GPT-4 have emerged as a potential game-changer in education, as a collaborative partner to assist teachers in learning design or to automatize decision-making processes. Despite known limitations, trained LLMs show promise in executing educational tasks. This study explores whether trained High-performance AI can facilitate teachers in the creation of boardgame-based learning units, by bridging their knowledge gap in game knowledge and game-based instructional skills. Using the GDBL ID model, the most comprehensive available instructional model for the creation of boardgame-based learning units, in this exploratory study we instructed Chat GPT to address two key phases of bGBL design: the choice of the game for the learning activity and the personalization of the game for constructive alignment and inclusion. Evaluation of the output by GBL experts highlights the potential of AI tools for bGBL

Keywords

Board Game-based learning; instructional design; GDBL ID; Personalization of games

1. Introduction

The advent of Artificial Intelligence (AI) has represented a revolution in different areas, including education. AI tools may redefine the role of the teacher and the entire education process, but their adoption is still at an embryonic stage and presents a number of open questions about prospects, risks and opportunities [1]. The use of Artificial Intelligence in the educational context (AIEd) has represented a growing interdisciplinary field since the 1970s to improve course design and expected student outcomes [2]. The aim is the creation of technological, innovative, and intelligent AI-powered systems to make learning personalized, engaging, and flexible. However, this process has introduced several opportunities and challenges for educational innovation, such as a move toward tailoring education to each student, challenging the traditional role of teachers, within an increasingly complex educational system. Indeed, the literature highlights the 'imperative change': AI is increasingly understood as an inevitable change to be adapted to meet the needs of a technology-based society. Second, especially in higher education, AI is decentralizing the figure of the teacher, spreading his or her role and authority among all the actors and tools involved in the learning process [3]. Artificial intelligence systems serving teaching and learning processes have distinctive traits that include grouping them into:

CEUR Workshop Proceedings (CEUR-WS.org)

- system supporting the process such as a platform or interactive learning environment to implement activities.
- tool that fulfills partial requests at different stages of the process like the use of assessment software.
- additional assistant that helps stakeholders better understand and direct the different features of the process [4].

Intelligent teaching assistants represent a frontier in teaching with AI to provide students and teachers with a range of functions to help and/or assist them in various activities of daily routine. Especially important for teachers is support in instructional design to promote effective and meaningful teaching [5]. For this reason, advanced AI systems such as GPT-4 are emerging as potential catalysts for change in the educational field. ChatGPT is a conversational assistant, launched by OpenAI in November 2022, that uses advanced natural language processing techniques to provide responses to received input. It can generate coherent, systematic, and information-rich responses [6]. Although there are recognized limitations on its use[7], there is a need to experiment with adherence to pedagogical frameworks [8] and methods such as Evolutionary algorithms [9] to develop quality educational resources [10]. For teachers, AI presents itself as an important resource since the use of intelligent tools elicits positive attitudes [11] and offers professional development: it improves teaching competence [12], instructional practices [13, 14], promotes self-reflection [15] and understanding of learning processes. Generative artificial intelligence tools, when integrated into a pedagogical framework for instructional design, can enhance the effectiveness of the teaching and learning process. By using these tools with a design approach, educators can create and deliver personalized and enriching educational pathways. Implementing the instructional design matrix introduces innovative methods to engage students, adapt content, and promote personalized learning [16]. However, most studies analyze the use of AI in education from a predominantly technological point of view, with few pedagogical lunges and rare moments of experimentation that do not allow for an in-depth understanding of the complex roles of AI in education and learning processes and the dimension from a design perspective. To fill these gaps, this research proposes an exploratory study to understand how to integrate AI in learning design. The purpose of the article is to investigate a design model for the use of AI in board game-based learning instructional design processes and use Chat GPT for choosing and customizing games, automizing some steps of the proposed model.

1.1. A working model for AI in board Game-Based Learning design

The ADDIE model [17] is one of the most well-known tools to design and manage a learning project. It is based on five main instructional phases which stand for Analysis, Design, Development, Implementation, and Evaluation (Table 1). A recent paper by Ch'ng [18] discusses the two distinct approaches that AI can have in instructional design: Assisted Intelligence and Autonomous Intelligence. The first helps humans to perform tasks faster and better; the second is based on automated decision-making processes, producing outcomes without human intervention. As Assisted Intelligence, AI can be used within the ADDIE framework to help designers in the analysis of needs using learning analytics, by using data to support the design, development, and implementation of instructional projects, and to help the evaluation process based on student performance. As Autonomous intelligence, AI can be trained to curate firsthand the learning journey and plan, generate personalized learning content, provide real-time assistance, and provide timely and automated assessment and evaluation. A recent systematic review by Yan et al. [19] has identified 53 different educational tasks that could potentially benefit from AI-based automation.

Within this context, in this study we discuss and provide preliminary evidence of the use of Autonomous Intelligence in the context of board-game based learning (bGBL) design. In recent years, gaming emerged as a valuable educational support, not only for its ability to engage and motivate students, but also to develop knowledge, skills, and competences [20–23]. Whereas

research in the last decades has mostly focused on digital games [24], board games have a longstanding tradition of use for educational purposes [25, 26] which is accompanied by growing empirical evidence [27]. Board games, far from being supplanted by their digital counterparts, are enjoying significant commercial success [28] driven by an ever-increasing number of original games. Board games are active, social, fun, and situated experiences which can be leveraged in a variety of educational scenarios [29, 30] to promote acquisition of disciplinary knowledge [27], key competences [31, 32], generate a sense of physical "togetherness" [33], can be employed in situation of social and economic disadvantage, and can be modified (or "modded"), to better align with disciplinary content [34–36]. Researchers have especially focused on the modification (also called "modding"), or personalization, of commercial board games [37] to better align with state curricula [38], educational contexts or specific learning goals [39–41] However, only recently researchers have started addressing the issue of guiding teachers in evaluating [42] and choosing [43] commercial board games for instructional purposes.

When harnessing the main difficulties in teacher's adoption of game-based learning (not limited to board games, but also including digital games), teachers report that they do not feel adequately prepared to include games within their curriculum [44, 45] and fear the increased workload [46]. The literature suggests that the main prerequisites for effective GBL adoption include the knowledge of games [29, 47], the ability to analyze games for their learning potential [48], and the ability to integrate games in the teaching and learning practice according to the desired learning goals [49]. bGBL requires teachers to carefully consider the underpinning pedagogic learning design: For instance, they can use either off-the-shelf games (games designed with a recreational purpose), educational/serious games (games that have been designed with a learning-first approach), modify, or personalize, existing games, or create new games *ad hoc* [50]. Furthermore, games can fill distinct roles within a learning unit according to the scenario for which they are used [29, 51]: to this aim, teachers must be able to gauge the relative costs and opportunities for each scenario.

Unfortunately, most teachers do not possess such competencies: according to a study by Persico et al. [52], Italian and English teachers' competence in the design of GBL activities is superficial and restricted to personal experience; this is reflected in a very limited adoption of games in Italian schools and higher education settings [53, 54]. Teachers seem to be familiar with board games that they played in their youth, such as *Monopoly, Risk, Naval Battle*, or traditional board games, such as *Chess*, but are not acquainted with modern board games [55]. To this aim, this study investigates whether AI systems can be used to help bridge the gap in teacher preparation by assisting them in the effective design and implementation of activities based on the use of board games. As to our current knowledge there is no available literature concerning the use of high-performance AI systems to support teachers in the design of bGBL, this explorative study aims to evaluate the potential of Chat GPT in automating specific tasks in bGBL design, specifically the choice of games according to the context, learning goals, and game goals and the personalization of games to obtain a better alignment between game and learning goals.

Within this context, the GDBL ID framework developed by Andreoletti and Tinterri [51] is currently the most in-depth reference for bGBL design; it has four defining features that help set it apart from existing instructional frameworks upon which the model iterates and improves:

- It is based on established pedagogical frameworks: the Technological Pedagogical and Content Knowledge (TPaCK) [56], the Play Curricular Reflection Discussion (PCaRD)[57], and the Inquiry, Communication, Construction and Expression (ICCE)[58].
- It integrates current best evidence from game studies, pedagogy, and cognitive science perspectives to game-based learning.
- It focuses on the different roles played by teachers in designing and applying game-based activities [59, 60].
- It focuses on the personalization of the game and learning experience, through both in-game and out-game changes to make help constructive alignment and promote accessibility [34, 37, 38, 61]

The GDBL ID framework is structured to facilitate the design of game-based teaching and learning activities, considering both instructional guidance as well as the need to progressively develop critical game literacy required to design and personalize game-based learning activities. It provides a step-by-step instructional guide structured on the five ADDIE steps (Table 1): starting from the definition and alignment of learning goals and game goals, assessment and evaluation criteria to the choice and personalization of the gaming. In particular, the development phase of the model includes three steps that are specific to bGBL [42]: game selection, game choice, and personalization (Table 1).

Table 1
Phases of the GDBL ID model (adapted from [50]).

ADDIE	Phase of GDBL model	Description	Specific to bGBL?
A(nalysis)	Analysis of needs	Determine students' characteristics and needs	No
D(esign)	Learning outcomes	Expected learning outcomes in terms of knowledge, skills, and attitudes	No
	Game goals	Definition of the goals attainable through gameplay	Yes
	Acceptable evidence	Quantitative and qualitative data that can be obtained to assess learning outcomes during and after gameplay	No
D(evelopment)	Game shortlist	Selection of a shortlist of potential games	Yes
	Game choice	Selection of the game based on alignment with design goals	Yes
	Game personalization	Modify either the game or the game experience for better alignment	Yes
	Instructional strategies	Define the lesson plan	No
I(mplement)	Implementation	Realize the activity in the classroom	No
E(valuate)	Evaluation and revision	Evaluate the impact of the activity and revise	No

1.2. Using Chat GPT to automate game choice and personalization.

In this exploratory study, we investigated the quality of Autonomous Intelligence design of Chat GPT in the design of bGBL. To this aim, we focused on two steps of the GDBL ID model that are specific to bGBL: the choice of a board game based on the alignment with the context, learning objectives and game goals, and the personalization of the game to improve this alignment.
we simplified the step of game selection in two ways:
First, the process is stepwise and does not allow backtracking from analysis to the choice of games, as in the original model ([50], p. 127).
Second, we limited game selection to the description of eight modern board games that were fed to the AI. The reason for this choice is that, since we referred to human experts to evaluate the quality of Chat GPT output, we had to make sure that they were competent in the games analyzed. Thus, the research questions for this study are the following:
1. Is Chat GPT able to identify salient features of modern board games?
2. Is Chat GPT able to choose appropriate board games according to context, learning objectives and game goals?
3. Is Chat GPT able to define adequate personalization strategies to better align existing games according to context, learning objectives, and game goals?
4. Is Chat GPT able to provide reasonable explanations for the choices made?

2. Methods

In this exploratory study, we analyzed Chat GPT's ability to automate two steps of the GDBL ID model specific to bGBL, namely Choice of the game and Personalization. To this aim, we asked a trained teacher to provide the Analysis of needs, learning goal, game goals and expected learning outcomes for a bGBL activity (Table 1). In parallel, the researchers curated a list of eight modern board games (*Carcassonne, Catan, Codenames, Concept, Dixit, Monopoly, Pandemic, Risk*) that the AI could choose from to develop the activity. The games were all commercially successful modern games who had previously been used by the expert evaluators in GBL activities. Successively, we wrote the prompts stepwise to allow Chat GPT to perform the different phases:

a) Describe the eight games according to Theme, Structure, Genre, and Main Mechanisms [43, 51].
b) Choose the most appropriate game for the learning task according to the given contest, learning goal, game goals, and expected learning outcomes, providing a rationale for the choice according to five pre-established criteria (Table 4).
c) Propose up to three in-game or around-game personalization [37, 51, 62] to better align the game experience with the goals of the activity.

For this study, we used Chat GPT with GPT-4 model (3rd of August 2023 version). We set the following parameters according to Sridhar et al. [8].

- temperature = 0.7 (standard)
- max_tokens = 5000
- top_p = 1 (standard)
- frequency_penalty = 0

At the moment of writing, there is clear evidence that the quality of the prompts is instrumental to improve the quality of the model answers [63, 64]; however, there is not yet an established methodology as to how to build effective prompts in AI-assisted and AI-autonomous instructional design. For this study, we manually designed prompts adhering to the recommendations of Bozkurt and Sharma [63], White et al. [65] and Liu [64]. We decided to use specific rather than generic prompts [64]. We first defined a persona and structured key ideas when they referred to specific aspects of the model [65]. We provided context for the answer and, for game selection and personalization steps, fed the model a few examples [66]. We present the prompts in tables according to Sridhar et al. [8]. We also asked Chat GPT to provide post hoc summary tables of the answers given (Tables 3, 6, and 7). Furthermore, to ensure that Chat GPT answers could be replicated, we ran the same prompts on two different machines, with separate Open AI account, to check for internal consistency of answers. ChatGPT has limited ability to evaluate the accuracy of the generated information, as it lacks the ability to assess the credibility of the data it was trained on [67]. Thus, for the evaluation of the answers, we asked three bGBL experts to independently evaluate AI answers according to an evaluation rubric based on four dimensions: Game description, choice of the game, personalization of the game, and rationale provided. We asked the experts to provide short qualitative feedback for each dimension.

3. Results

First, the AI was asked to provide a short description of the games pre-selected according to their main characteristics (Table 2). This step was taken to ensure that the AI was able to extrapolate the salient features of the games in line with the GDBL ID model ([51], p. 129).

Table 2
The table shows the user message specifying the context for the game descriptions to be generated.

[USER] You are a curricular development expert system focused on board Game-based learning (bGBL). bGBL is a learning strategy aimed at promoting knowledge through the use of board games in the classroom. Your first task will be to describe five board games according to the following characteristics: 1) Theme [147 characters] 2) Structure. [219 characters] 3) Genre. [257 characters] 4) Main mechanism (or mechanic). [211 characters] The user will provide you with a few examples, then will give you the names of the board games you must describe.

[4 sample descriptions (315 characters)]

Chat GPT provided descriptions for the games summarized in Table 3:

Table 3
Chat GPT-generated table summarizing the description for each game.

Game	Theme	Structure	Genre	Main mechanism
Dixit	Abstract, Imagination	Multiplayer	Party Game, Family	Storytelling, Voting
Pandemic	Global Epidemic	Cooperative	Strategy, Family	Hand Management, Set Collection, Point-to-point Movement
Codenames	Espionage, Word Association	Two-Team Game	Party Game, Family	Word Association, Deduction, Team-based play
Concept	Abstract; Communication	Multiplayer	Party Game, Family	Deduction, Pictorial Communication
Carcassonne	Medieval Southern France	Multiplayer	Family, Eurogame	Tile Placement, Area Control
Monopoly	Real Estate, Capitalism	Multiplayer	Family, Strategy	Roll and Move, Trading, Property Management
Risk	Global Domination, Warfare	Multiplayer	Strategy, Wargame	Area Control, Dice Rolling, Player Elimination
Catan	Island Settlement	Multiplayer	Family, Strategy, Eurogame	Trading, Hand Management, Route Building

All experts independently rated Chat GPT's ability to describe the main feature of the games as "Good". Experts agreed that descriptions were generally accurate, with a few imprecisions. One expert noted that "The themes identified were sometimes very articulate (Medieval Southern France) and sometimes slightly inaccurate (Codenames is mostly abstract)." Second, we asked the AI to choose among the described games which fitted better with the provided instructional context, learning goals, and game goals (Table 4).

Table 4

The table shows the user message specifying the context for the choice of the game and, in blue, Chat GPT's response.

[USER]The user will provide you the context, learning goals and game goals of a bGBL learning unit. You will evaluate which of the games you described (Dixit, Pandemic, Codenames, Concept, Carcassonne, Monopoly, Risk) is more suited to achieve the learning and game goals. You will argue this decision based on five key indicators. 1) Accessibility 2) Setting (time required to set up and play the game) 3) Content (whether the content/thematic knowledge is used to address game challenges or is merely accessory 4) Learning Curve 5) Opportunities to assess success standards provided internally by the game.

[Translated from Italian] Class III secondary school. The class consists of 24 pupils and is rather heterogeneous in terms of interest, application in study and participation in teaching activities. The class group shows lively but essentially responsible behavior. The levels of basic skills are heterogeneous. [362 characters]. Pupils with Special Educational Needs: 1 student with DSA - dyslexia, reading and writing difficulties. Benefit from special teaching materials and the use of computers to overcome the barriers of dyslexia. 1 student with ADHD - Attention Deficit/Hyperactivity Disorder, difficulty maintaining attention for prolonged periods; benefits from short learning sessions and frequent breaks.

LEARNING GOAL: The learner interacts effectively in a variety of communicative situations, through dialogic modes that are always respectful of the ideas of others. Uses metalinguistic knowledge to understand the meanings of texts more precisely.

EXPECTED LEARNING OUTCOMES:

Intervene in a conversation or discussion, class or group, with relevance and coherence, respecting timing and turns to speak and making a positive personal contribution.

Use one's knowledge of text types to adopt functional strategies for understanding while listening.

GAME OBJECTIVES:

Use metalinguistic skills to understand context and communication skills to interact effectively.

Take on different roles to achieve specific goals in the game.

Use active listening and constructive dialogue strategies to counter opposing arguments.

Use dialogue and metalinguistic skills to reach a win-win agreement.

Use communication and metalinguistic skills to solve problems in a game environment that simulates a real-world situation.

[CHAT GPT]

Codenames:

Accessibility: High. Word-based game, suitable for language-focused objectives.

Setting: Quick setup, games last about 15-30 minutes.

Content: Central. The game revolves around understanding and communicating word associations.

Learning Curve: Low. Simple rules, easy to grasp.

Opportunities for Assessment: High. The game provides immediate feedback through correct and incorrect guesses. [1152 characters]

While Concept is also a strong contender due to its emphasis on non-verbal communication, the specificity of the learning goals around metalinguistic knowledge and active dialogue make Codenames a slightly better fit for this particular context.

The three experts agreed that *Codenames* was the best choice among the available games in terms of alignment for the context. One expert noted that *Codenames*, despite being a game based on words, was playable also by student with dyslexia. Another noted that "the game's short duration and setup, coupled with its "tight" gameplay, (...) fits well with the characteristics of the class and

the presence of an ADHD student". Furthermore, the experts agreed on the quality of the choice considering learning goals. "Codenames is a sensible choice for the goal of developing metalinguistic competence, as it allows reasoning on secondary or metaphorical meanings of words. It also has different roles and teamplay which help improve effective communication between students" wrote one expert. Finally, all experts agreed that the chosen game was the best choice also considering the game goals defined by the teacher. One expert adds: "Codenames fits all the game goals of the units, with the partial exception of the condition "simulates a real environment" as the game is mostly abstract. Another sensible choice would have been Pandemic, which was however less focused on metalinguistic communication". The last step that we asked the AI to automatize was the choice of personalizations that could help a tighter alignment between the goals of the learning unit and the game of choice. A few examples of in-game and around-game personalization [51] were given to the AI as part of the input (Table 5).

Table 5
The table shows the user message specifying the context for the game personalization.

[USER] You can personalize (modify) the game you selected to better align with the needs analysis, learning objectives and game goals. Personalization can be either in game or around game. In game personalization implies altering the actions of the player via changes to the game rules and/or pieces. Around game personalization acts on the game experience, including interaction among players, between players and the teacher as well as the context where the game activity takes place. The user will provide a few examples of in game and around game personalization [669 characters].

The AI suggested three personalizations, two around-game and one in-game, summarized in Table 6.

Table 6
Chat GPT-generated table summarizing the game personalization proposed.

Personalization Type	Description	Rationale
Around Game	Play in couples for peer learning and support	Promotes collaboration and peer learning. Supports students with special education needs.
Around Game	Insert moments of reflection after each round/game for deeper understanding and discussion	Encourages deeper understanding, discussion, and alignment with metalinguistic objectives.
In Game	Modify luck/skill balance by introducing "discussion tokens" to promote more dialogue and metalinguistic understanding	Promotes more dialogue, emphasizes metalinguistic understanding, and offers an additional strategy layer.

The decision to play in couples was motivated by the AI system as follows:
- It allows stronger students to support peers with challenges, fostering collaboration and peer learning.
- The student with ADHD, who benefits from short learning sessions, can switch roles with their partner, ensuring active engagement without being overwhelmed.
- The student with DSA (dyslexia) can have a partner to help with reading, reducing the barrier and allowing them to focus on the metalinguistic and communicative aspects of the game.

The second personalization, the inclusion of moments of reflection after each round, is aimed at having students discuss and reflect on in-game events, such as their choice of clues and the rationale behind them, how they interpreted the clues given by the spymaster, and the

metalinguistic challenges they faced. According to the AI, this personalization "aligns well with the learning goals of understanding and using metalinguistic knowledge and effectively interacting in communicative situations."

Finally, Chat GPT proposed a new in-game rule: the introduction of "discussion tokens".

"Each team gets a limited number of these tokens at the start. Using a token allows the team an additional 30 seconds to discuss and deliberate on a particularly challenging clue."

According to the AI, this rule "promotes more dialogue and emphasizes the importance of metalinguistic understanding and communication over the rush of the regular timer. It also provides an additional layer of strategy, as teams must decide when it's most beneficial to use these tokens".

All three experts agreed that the proposed personalizations would be very fitting given the class context. One expert notes: "Playing in couples is potentially beneficial for students with special needs and allows them to play the game together with their peers. The classroom shows interest in discussion when appropriately stimulated, therefore the other personalization could improve cognitive, social, and emotional engagement with the learning activity". The experts judge that the personalizations are a good fit with the learning objectives, although one notes: "The rationale for ordered discussion is, in my opinion, not treated adequately. The choice of tokens for group discussion is interesting but this could make the order of player intervention chaotic. Maybe defining a set number of tokens for each player could mitigate this issue". Another expert notes that "the personalizations provide opportunities for students to intervene and discuss (...) however they barely address the second learning goal related to knowledge of different types of text". The experts also convene that the personalizations align significantly with the game goals, and all three experts agree that they are sufficiently easy and could realistically be implemented in the classroom setting.

The experts were asked to rate the quality of the rationale provided by the AI for the choices taken. All experts agreed that the rationale for the game selected was good but could be better motivated. All three note how synthetic the explanation is. One expert suggests that the prompt could specify a point-per-point comparison with the excluded games. The second highlights that the discussion concerning "accessibility" did not mention students with special needs and how the game suited them; "furthermore, the AI seem to interpret differently the cue-giving in Codenames compared to Concept, whereas it is basically the same mechanism". Finally, the third expert hypothesizes that a more specific request prompt would have probably returned a more satisfying explanation. Conversely, the experts agree that the rationale for the excluded game is very convincing (Table 7).

Table 7

Chat GPT-generated table summarizing the game inclusion/exclusion rationale.

Game	Criteria for Inclusion	Criteria for Exclusion
Codenames	Emphasizes word associations, cooperative nature, and active listening. Aligns with metalinguistic and communicative objectives.	N/A (Selected game)
Dixit	Promotes creativity, storytelling, and interpretation.	Abstract nature may not directly address metalinguistic knowledge or focused communication skills.
Pandemic	Fosters teamwork, strategic thinking, and collective decision-making.	Theme and mechanics more about disease control, less on metalinguistic understanding or structured conversation.

Carcassonne	Emphasizes spatial reasoning, strategy, and area control.	More about building territories than active dialogue or metalinguistic challenges.
Monopoly	Involves negotiation, property management, and strategic decision-making.	Can be lengthy and might not focus on metalinguistic and communicative skills targeted.
Risk	Promotes global strategy, player interactions, and battles	Focused on warfare and global strategy, less on active communication.
Concept	Emphasizes non-verbal communication and understanding of universal icons.	While focusing on non-verbal communication, it might be less direct in addressing metalinguistic understanding compared to Codenames.

One expert notes that "The AI does an impressive job of identifying differences between Dixit and Codenames, considering that both are language- and cue- based game". Finally, two experts considered that the rationale given for the personalizations is convincing but should be better explained and more user-friendly if addressed to non-experts. The expert evaluations are summarized in Table 8.

Table 8

Summary of the expert evaluation for Chat GPT performance.

Dimension	Indicator(s)	Excellent	Good	Average	Poor
Game description	The description of the proposed games is precise in terms of Theme, Structure, Genre, and Mechanism		◎◎◎		
Game choice	The chosen game fits the class context	◎◎◎			
	The chosen game fits the learning goals of the unit	◎◎◎			
	The chosen game fits the game goals	◎◎◎			
Game personalization	The proposed personalization(s) fits the class context	◎◎◎			
	The proposed personalization(s) fit the learning goals of the unit		◎◎◎		
	The proposed personalization(s) fit the game goals	◎◎	◎		
	The proposed personalization(s) are easy to implement	◎◎◎			
Rationale	The rationale provided for the chosen game is convincing		◎◎◎		
	The rationale provided for the excluded game is convincing	◎◎	◎		
	The rationale provided for the personalization(s) is convincing	◎	◎◎		

4. Discussion

In this study, we started an exploration of the potential of Chat GPT as a support for teachers in the design of bGBL learning units. The literature highlights how, despite a generalized trust for the learning potential and effectiveness of GBL[53, 54], games are seldom used in the school and other formal learning contexts. This seems due, at least in part, to lack of game knowledge and lack of GBL competence by teachers. Our hypothesis is that high-performance AI models can help bridge the gaps in teacher preparation, not as a substitute for the teachers' creativity and professional competence but to help them overcome their fears due to lack of confidence with the medium and jumpstart the use of games in the classroom by supporting the design, development, implementation, and evaluation of bGBL activities. Our preliminary results indicate that Chat GPT has the potential to provide effective support in several steps of bGBL design, notably the choice of games according to the students' characteristics and needs and the specific learning and game goals identified, as well as the personalization of in-game and around-game aspects to achieve a better constructive alignment between the existing game and learning goals. The experts were largely in agreement that the AI was able to extract the salient characteristics of the games proposed (RQ1), selected the game that fit best for the proposed activity (RQ2), suggested sensible, realistic, and useful personalizations (RQ3), and gave reasonable justification for the choices operated (RQ4). However, according to the nature of an exploratory study, the results are subject to several limitations: In the first place, the analysis is limited to a single learning unit; it can be argued that Chat GPT's ability to provide a sensible choice could be due to a particularly good match between learning settings and the game of choice. Providing a wider range of learning contexts, specific learning and game goals falls beyond the scope of this exploratory study: however, this issue should be explored in further studies to ensure reproducibility. Second, we restricted the task of AI by only providing the choice of the game from within a limited list. This effectively prevented us to test Chat GPT ability to select from a wider range of options and could also have facilitated the game analysis step, as the games were rather heterogeneous in their nature and mechanics. The model was very good at discriminating and motivating the choice between games based on linguistic cues (Codenames, Concept) and metalinguistic ability (Codenames, Concept). However, further, finer-grained analysis between games with similar mechanisms would be more apt at stressing Chat GPT's abilities further. Third, due to the experimental setup we could only test AI-autonomous creation, whereas AI-assisted creation is likely to be at least as interesting an avenue to develop bGBL units and promote teacher upskilling in game knowledge and GBL design. Fourth, the setup allowed no comparison between human-made and AI-made decisions. Future studies could address this issue by blind evaluation of human-developed and AI-developed GBL units to better understand the model potential. Fifth, the experts were aware of evaluating AI-created decisions, as they were provided the entire conversation. This, depending on each expert's views and beliefs concerning AI, could have influenced their judgements. Sixth, we could not test empirically the quality of AI-generated choices. Future research could address this issue by testing AI-designed bGBL units in live environments and addressing their impact. Seventh, we only tested Chat GPT on the GPT-4 model; to ensure internal consistency, we ran the same prompt on a different account and obtained largely overlapping answers (data not shown). However, we are aware that the results obtained with different models, such as GPT-3.5 or Google Bard, could have led to significantly different results. GPT-4 being a premium, paid-for tool, its implementation in educational contexts might be limited. Future research will need to address this issue by focusing on comparative analysis and prompting techniques.

5. Conclusion

Despite the limitations of an exploratory study, to our knowledge this is the first research addressing the potential of high-performance AI models in game-based learning design, and the early returns are, to say the least, very promising. Future research will clarify whether high-

performance AI models can be effectively leveraged to finally allow the diffusion of game-based learning, overcoming the difficulties that historically prevented the adoption of this extremely promising, but severely underutilized, learning methodology.

Author contributions

Conceptualization, A.T.; methodology, A.T, F.P. and M. di P.; investigation, A.T.; evaluation, A.T., G.V., and F.P.; writing—original draft preparation, A.T., M. di P.; writing—review and editing, A.T., M. di P., F.P., G.V., and A.D.; supervision, A.D. All authors have read and agreed to the published version of the manuscript.

References

1. E. Dimitriadou, A. Lanitis, A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms. Smart Learn. Environ. 10, 2023. https://doi.org/10.1186/s40561-023-00231-3.
2. E. Southgate, K. Blackmore, S. Pieschl, S. Grimes, J. McGuire, K. Smithers, Artificial intelligence and emerging technologies in schools: Research report. University of Newcastle, Newcastle, NSW, 2019.
3. M. Bearman, J. Ryan, R. Ajjawi, Discourses of artificial intelligence in higher education: a critical literature review. High. Educ. 86, 2023, pp. 369–385. https://doi.org/10.1007/s10734-022-00937-2.
4. W. Xu, F. Ouyang, A systematic review of AI role in the educational system based on a proposed conceptual framework. Educ. Inf. Technol. 27, 2022. pp. 4195–4223. https://doi.org/10.1007/s10639-021-10774-y.
5. X. Zhang, J. Geng, Y. Chen, S. Hu, T. Huang, I-assistant: An Intelligent Teaching Assistant System for Classroom Teaching. In: 2023 IEEE 12th International Conference on Educational and Information Technology (ICEIT), 2023, pp. 1–5. https://doi.org/10.1109/ICEIT57125.2023.10107774.
6. T. Adiguzel, H. Kaya, F. Cansu, Revolutionizing education with AI: Exploring the transformative potential of ChatGPT, Contemp. Educ. Technol. 15, 2023, ep429. https://doi.org/10.30935/cedtech/13152.
7. D. Baidoo-Anu, L. Owusu Ansah, Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning, SSRN, 2023. https://doi.org/10.2139/ssrn.4337484.
8. P. Sridhar, A. Doyle, A. Agarwal, C. Bogart, J. Savelka, M. Sakr, Harnessing LLMs in Curricular Design: Using GPT-4 to Support Authoring of Learning Objectives, ARXIV, 2023. https://doi.org/10.48550/arXiv.2306.17459.
9. S. Elsayed, Towards Mitigating ChatGPT's Negative Impact on Education: Optimizing Question Design through Bloom's Taxonomy, ARXIV, 2023. https://doi.org/10.48550/arXiv.2304.08176.
10. C.K. Lo, What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature, Educ. Sci. 13, 2023. https://doi.org/10.3390/educsci13040410.
11. S.A.M. Aldosari, The Future of Higher Education in the Light of Artificial Intelligence Transformations, Int. J. High. Educ. 9, 145, 2020. https://doi.org/10.5430/ijhe.v9n3p145.
12. A. Jaiswal, C.J. Arun, Potential of Artificial Intelligence for Transformation of the Education System in India, Int. J. Educ. Dev. Using Inf. Commun. Technol 17, 2021, pp. 142–158.
13. K.D.H. Gunawan, L. Liliasari, I. Kaniawati, W. Setiawan, Implementation of Competency Enhancement Program for Science Teachers Assisted by Artificial Intelligence in Designing HOTS-based Integrated Science Learning. J. Penelit. Dan Pembelajaran IPA 7, 2021, pp. 55–65. https://doi.org/10.30870/jppi.v7i1.8655.

14. J. Hu, Teaching Evaluation System by use of Machine Learning and Artificial Intelligence Methods, Int. J. Emerg. Technol. Learn. IJET 16, 2021, pp. 87–101. https://doi.org/10.3991/ijet.v16i05.20299.
15. N.L.S. Aldeman, K.M. de Sá Urtiga Aita, V.P. Machado, L.C.D. da Mata Sousa, A.G.B. Coelho, A.S. da Silva, A.P. da Silva Mendes, F.J. de Oliveira Neres, S.J.H. do Monte, Smartpathk: a platform for teaching glomerulopathies using machine learning, BMC Med. Educ. 21, 2021. https://doi.org/10.1186/s12909-021-02680-1.
16. L.I. Ruiz-Rojas, P. Acosta-Vargas, J. De-Moreta-Llovet, M. Gonzalez-Rodriguez, Empowering Education with Generative Artificial Intelligence Tools: Approach with an Instructional Design Matrix, Sustainability 15, 2023. https://doi.org/10.3390/su151511524.
17. M. Molenda, In search of the elusive ADDIE model. Perform. Improv. 42, 2003, pp. 34–36. https://doi.org/10.1002/pfi.4930420508.
18. L.K. Ch'ng, How AI Makes its Mark on Instructional Design, 2023. https://doi.org/10.5281/ZENODO.8188576.
19. L. Yan, L. Sha, L. Zhao, Y. Li, R. Martínez-Maldonado, G. Chen, X. Li, Y. Jin, D. Gašević, Practical and ethical challenges of large language models in education: A systematic scoping review, Br. J. Educ. Technol, 2023. https://doi.org/10.1111/bjet.13370.
20. S. De Freitas, G. Rebolledo-Mendez, F. Liarokapis, G. Magoulas, A. Poulovassilis, Learning as immersive experiences: Using the four-dimensional framework for designing and evaluating immersive learning experiences in a virtual world, Br. J. Educ. Technol 41, 2010, pp. 69–85. https://doi.org/10.1111/j.1467-8535.2009.01024.x.
21. J.P. Gee, What Video Games Have to Teach Us About Learning and Literacy, St. Martin's Griffin, 2003.
22. J.L. Plass, B.D. Homer, E.O. Hayward, J. Frye, T.-T. Huang, M. Biles, M. Stein, K. Perlin, The Effect of Learning Mechanics Design on Learning Outcomes in a Computer-Based Geometry Game, In: E-Learning and Games for Training, Education, Health and Sports, Springer, Berlin, Heidelberg, 2012, pp. 65–71. https://doi.org/10.1007/978-3-642-33466-5_7.
23. V. Gashaj, L.C. Dapp, D. Trninic, C.M. Roebers, The effect of video games, exergames and board games on executive functions in kindergarten and 2nd grade: An explorative longitudinal study, Trends Neurosci. Educ. 25, 2021. https://doi.org/10.1016/j.tine.2021.100162.
24. J.L. Plass, R.E. Mayer, B.D Homer (Eds.), Handbook of Game-Based Learning, MIT Press, Cambridge, MA, USA, 2020.
25. T. Donovan, It's All a Game: The History of Board Games from Monopoly to Settlers of Catan, Macmillan, London, UK, 2017.
26. F. Gobet, J. Retschitzki, A. de Voogt, Moves in Mind: The Psychology of Board Games, Psychology Press, London, UK, 2004.
27. R.Y. Bayeck, Examining Board Gameplay and Learning: A Multidisciplinary Review of Recent Research, SAGE Journals, 2020. https://doi.org/10.25384/SAGE.c.4943763.
28. ResearchAndMarkets.com, Board Games Market: Global Outlook and Forecast 2023-2028, 2023, URL: https://www.researchandmarkets.com/reports/5750177/board-games-market-global-outlook-forecast.
29. T. Hanghøj, Game-Based Teaching: Practices, Roles, and Pedagogies, New Pedagogical Approaches in Game Enhanced Learning: Curriculum Integration, IGI Global, 2013, pp. 81–101. https://doi.org/10.4018/978-1-4666-3950-8.ch005.
30. K. Salen, Designing a Place Called School: A Case Study of the Public School Quest to Learn, She Ji J. Des. Econ. Innov. 3, 2017, pp. 51–64. https://doi.org/10.1016/j.sheji.2017.08.002.
31. M. Berland, S. Duncan, Computational Thinking in the Wild: Uncovering Complex Collaborative Thinking through Gameplay, Educ. Technol, 56, 2016, pp. 29–35.
32. M. Berland, V.R. Lee, Collaborative Strategic Board Games as a Site for Distributed Computational Thinking, Int. J. Game-Based Learn, 1, 2011, pp. 65–81. https://doi.org/10.4018/ijgbl.2011040105.
33. G. Calleja, Unboxed: Board Game Experience and Design, MIT Press, Cambridge, MA, 2022.
34. E. Castronova, I. Knowles, A Model of Climate Policy Using Board Game Mechanics, Int. J. Serious Games, 2, 2015. https://doi.org/10.17083/ijsg.v2i3.77.

35. M. Akcaoglu, Learning problem-solving through making games at the game design and learning summer program, Educ. Technol. Res. Dev, 62, 2014, pp. 583–600. https://doi.org/10.1007/s11423-014-9347-4.
36. M. Akcaoglu, A. Gutierrez de Blume, P. Sonnleitner, C. Hodges, Game Design as a Complex Problem Solving Process, in: R. Zheng & M. Gardner (Eds.), Handbook of Research on Serious Games for Educational Applications, IGI Global, 2016. https://doi.org/10.4018/978-1-5225-0513-6.ch010.
37. D. Abbott, Modding Tabletop Games for Education, in: M. Gentile, M. Allegra, H. Söbke, (Eds.) Games and Learning Alliance, Springer International Publishing, Cham, 2019, pp. 318–329. https://doi.org/10.1007/978-3-030-11548-7_30.
38. N.B. Sardone, Modifying Board Games in Alignment with State Standards to Develop the Geographic Literacy of Elementary Level Learners, Soc. Stud. 0, 2022, pp. 1–11. https://doi.org/10.1080/00377996.2022.2114403.
39. S. Nicholson, Making Gameplay Matter: Designing Modern Educational Tabletop Games, Knowl. Quest. 40, 2011, pp. 60–65.
40. A.R. Denham, Improving the Design of a Learning Game Through Intrinsic Integration and Playtesting, Technol. Knowl. Learn. 21, 2016, pp. 175–194. https://doi.org/10.1007/s10758-016-9280-1.
41. M. Sousa, Modern Serious Board Games: modding games to teach and train civil engineering students, 2020 IEEE Global Engineering Education Conference (EDUCON), 2020, pp. 197–201. https://doi.org/10.1109/EDUCON45650.2020.9125261.
42. M. Sousa, Mastering Modern Board Game Design to Build New Learning Experiences: the MBGTOTEACH Framework, Int. J. Games Soc. Impact. 1, 2023, pp. 68–93.
43. S.P. Greenhalgh, M.J. Koehler, L.O. Boltz, The Fun of Its Parts: Design and Player Reception of Educational Board Games, CITE Journal, 3, 2019.
44. Y. Allsop, E. Yildirim, M. Screpanti, Teachers' Beliefs About Game Based Learning: A Comparative Study of Pedagogy, Curriculum and Practice in Italy, Turkey and the UK, 7th ECGBL Conference Proceedings, 2013.
45. L.M. Takeuchi, S. Vaala, Level up Learning: A National Survey on Teaching with Digital Games, Joan Ganz Cooney Center at Sesame Workshop, 2014.
46. T. Nieland, A. Fehrenbach, M. Marowsky, M. Burfeind, The Teacher-Centered Perspective on Digital Game-Based Learning, in: C. Aprea, D. Ifenthaler, (Eds.) Game-based Learning Across the Disciplines, Springer International Publishing, Cham, 2021, pp. 341–362. https://doi.org/10.1007/978-3-030-75142-5_15.
47. Y.K. Baek, What hinders teachers in using computer and video games in the classroom? Exploring factors inhibiting the uptake of computer and video games, Cyberpsychology Behav. Impact Internet Multimed. Virtual Real. Behav. Soc 11, 2008, pp. 665–671. https://doi.org/10.1089/cpb.2008.0127.
48. E. Koh, Y.G. Kin, B. Wadhwa, J. Lim, Teacher Perceptions of Games in Singapore Schools, Simul. Gaming. 43, 2012, pp. 51–66. https://doi.org/10.1177/1046878111401839.
49. A.D. Ritzhaupt, E. Gunter, Survey of commercial off-the-shelf video games, benefits and barriers in formal educational settings, AECT Conference, 2010.
50. N. Whitton, The Place of Game-Based Learning in an Age of Austerity, Electron. J. E-Learn. 10, 2012, pp. 249–256.
51. M. Andreoletti, A. Tinterri, Apprendere con i giochi. Esperienze di progettazione ludica. Carocci Editore, Roma, 2023.
52. D. Persico, M. Passarelli, F. Pozzi, J. Earp, F. Dagnino, F. Manganello, Meeting players where they are: Digital games and learning ecologies, Br. J. Educ. Technol. 50, 2019. https://doi.org/10.1111/bjet.12777.
53. Y. Allsop, , J. Jessel, Teachers' Experience and Reflections on Game-Based Learning in the Primary Classroom: Views from England and Italy, Int. J. Game-Based Learn. 5, 2015, pp. 1–17. https://doi.org/10.4018/ijgbl.2015010101.
54. F.F. Loperfido, A. Dipace, A. Scarinci,: TO PLAY OR NOT TO PLAY? A CASE STUDY OF TEACHERS' CONFIDENCE AND PERCEPTION WITH REGARD TO DIGITAL GAMES AT

SCHOOL, Ital. J. Educ. Technol. 27, 2019, pp. 121–138. https://doi.org/10.17471/2499-4324/1062.

55. M. Cantoia, A. Clegg, A. Tinterri, Training Teachers to Design Game-Based Learning Activities: Evidence from a Pilot Project, Computers in the Schools, 2023. https://doi.org/10.1080/07380569.2023.2271462.

56. P. Mishra, M.J. Koehler, Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge, Teach. Coll. Rec. 108, 2006, pp. 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x.

57. A. Foster, Assessing Learning Games for School Content: The TPACK-PCaRD Framework and Methodology, in: D. Ifenthaler, D. Eseryel, X. Ge, (Eds.) Assessment in Game-Based Learning: Foundations, Innovations, and Perspectives, Springer, New York, NY, 2012, pp. 201–215. https://doi.org/10.1007/978-1-4614-3546-4_11.

58. A. Foster, M. Shah, The ICCE Framework: Framing Learning Experiences Afforded by Games, J. Educ. Comput. Res. 51, 2015, pp. 369–395. https://doi.org/10.2190/EC.51.4.a.

59. T. Hanghøj, C. Brund, TEACHER ROLES AND POSITIONINGS IN RELATION TO EDUCATIONAL GAMES, 2011. URL: https://doi.org/10.2307/jj.608141.9.

60. M. Kangas, A. Koskinen, L. Krokfors, A qualitative literature review of educational games in the classroom: the teacher's pedagogical activities, Teach. Teach. 23, 2017, pp. 451–470. https://doi.org/10.1080/13540602.2016.1206523.

61. M. Sousa, Serious board games: modding existing games for collaborative ideation processes, Int. J. Serious Games. 8, 2021, pp. 129–146. https://doi.org/10.17083/ijsg.v8i2.405.

62. D. Abbott, Intentional Learning Design for Educational Games: A Workflow Supporting Novices and Experts, Learn. User Exp. Res., 2020.

63. A. Bozkurt, R.C. Sharma, Generative AI and Prompt Engineering: The Art of Whispering to Let the Genie Out of the Algorithmic World, Asian J. Distance Educ. 18, 2023, pp. 1-7. https://doi.org/10.5281/zenodo.8174941.

64. L. Liu, Analyzing the Text Contents Produced by ChatGPT: Prompts, Feature-Components in Responses, and a Predictive Model, J. Educ. Technol. Dev. Exch. JETDE, 16, 2023, pp. 49–70. https://doi.org/10.18785/jetde.1601.03.

65. J. White, Q. Fu, S. Hays, M. Sandborn, C. Olea, H. Gilbert, A. Elnashar, J. Spencer-Smith, D.C. Schmidt, A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT, ARXIV, 2023. http://arxiv.org/abs/2302.11382,

66. P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, G. Neubig, Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing, ACM Comput. Surv. 55, 2023. https://doi.org/10.1145/3560815.

67. M. Farrokhnia, S.K. Banihashem, O. Noroozi, A. Wals, A SWOT analysis of ChatGPT: Implications for educational practice and research. Innov. Educ. Teach. Int. 0, 2023, pp. 1–15. https://doi.org/10.1080/14703297.2023.2195846.