

Integrating GPT as an Assistant for Low-Cost Virtual Reality Escape-Room Games

Alan Rychert, *Department of Computer Science and Engineering, Universidad Nacional del Sur, Bahía Blanca, 8000, Argentina*

María Luján Ganuza ^{id} and Matias Nicolás Selzer ^{id}, *Institute for Computer Science and Engineering (UNS-CONICET), and VyGLab Research Laboratory (UNS-CICPBA), Department of Computer Science and Engineering, Universidad Nacional del Sur, Bahía Blanca, 8000, Argentina*

This work explores the integration of generative pretrained transformer (GPT), an AI language model developed by OpenAI, as an assistant in low-cost virtual escape games. The study focuses on the synergy between virtual reality (VR) and GPT, aiming to evaluate its performance in helping solve logical challenges within a specific context in the virtual environment while acting as a personalized assistant through voice interaction. The findings from user evaluations revealed both positive perceptions and limitations of GPT in addressing highly complex challenges, indicating its potential as a valuable tool for providing assistance and guidance in problem-solving situations. The study also identified areas for future improvement, including adjusting the difficulty of puzzles and enhancing GPT's contextual understanding. Overall, the research sheds light on the opportunities and challenges of integrating AI language models such as GPT in virtual gaming environments, offering insights for further advancements in this field.

In an increasingly digitalized world, the convergence of virtual reality (VR), augmented reality (AR), and artificial intelligence (AI) has opened new frontiers of innovation. VR, a technology once considered unattainable, is now a tangible part of our society. Its ability to immerse people in 3-D environments akin to reality has piqued the interest of a broad audience, ranging from gaming enthusiasts to professionals in fields such as medicine and education.

In recent years, there have been remarkable advancements in AI, sparking a revolution across multiple industries and fundamentally reshaping our interactions with technology. One of the most significant advancements in this field is the development of large language models (LLMs), such as generative pretrained transformer (GPT). GPT is a type of language model that uses deep learning techniques to generate

human-like text based on the patterns it has learned from vast amounts of data. It is trained on a massive corpus of text from the Internet, allowing it to understand and generate natural language with remarkable fluency and coherence. ChatGPT, on the other hand, is a conversational AI system built on top of GPT technology. It is designed to engage in natural language interactions, understanding context and providing relevant, coherent responses to user queries or prompts.¹⁰

Virtual reality escape rooms have gained significant popularity in recent years, offering immersive and engaging experiences across domains such as education, emotion elicitation, and entertainment. These interactive adventures challenge players to solve logic-based puzzles and riddles through collaboration and contextual understanding, making them an ideal testbed for evaluating the capabilities of AI language models like GPT. The integration of GPT as a personalized assistant in VR escape rooms has the potential to enhance player engagement by providing real-time guidance, adapting to player decisions, and fostering a dynamic problem-solving partnership. This innovative setup not only pushes the boundaries of traditional

0272-1716 © 2024 IEEE
Digital Object Identifier 10.1109/MCG.2024.3426314
Date of publication 11 July 2024; date of current version 20 August 2024.

gameplay but also showcases the potential of GPT to improve gaming experiences by using its contextual understanding, puzzle-solving abilities, and interactive collaboration with users in immersive virtual environments.

This research study focuses on investigating the integration of GPT as an assistant in low-cost VR escape-room games. The main research question focuses on understanding how GPT enhances user experience in virtual escape games by assisting users in solving complex challenges. The study aims to evaluate the effectiveness of GPT as a companion in the virtual escape room environment, the usability of the system, and the difficulty of the puzzles. By analyzing user perceptions and feedback, the research seeks to identify areas for improvement in optimizing the user experience and enhancing GPT's capabilities in providing clear and contextually relevant responses. The context of the study involves exploring the potential of AI-powered assistants such as GPT to enhance interactive experiences in VR environments, particularly in the context of puzzle-solving and cognitive challenges within escape room scenarios.

This article demonstrates the innovative application of GPT as a dynamic assistant in low-cost VR escape rooms. By using GPT's advanced natural language processing (NLP) capabilities, we have created an interactive environment where AI-adaptive dialogs enhance the player's experience and engagement. This approach not only transforms the traditional escape room model into a more interactive and personalized experience but also offers a cost-effective solution for educational purposes. By integrating low-cost VR hardware with sophisticated AI, our study highlights the potential for scalable, immersive learning tools that can be widely accessible. This work contributes to the growing body of research demonstrating the practical applications of AI in VR settings, paving the way for future developments in both educational technologies and interactive entertainment systems.

RELATED WORK

This work integrates several related areas: low-cost VR systems, escape room games, NLP with models such as GPT, and interactive problem-solving scenarios.

Low-Cost VR

VR is a technology that allows users to interact with a computer-generated 3-D environment, providing an immersive experience that simulates physical presence within that environment. This simulation is

achieved through specialized hardware devices, such as VR headsets equipped with sensors, which capture user movements and translate them into corresponding actions within the virtual environment.¹³ This technology has various applications in fields such as entertainment, education, medicine, design, and architecture, retail, as well as military and emergency training. For instance, in education, VR enables the creation of realistic simulations, while in medicine, it is used for exposure therapies and rehabilitation purposes.⁶

Low-cost VR refers to VR experiences that can be accessed using less expensive hardware, such as smartphones and affordable headsets such as Google Cardboard. These devices allow users to experience VR without the need to invest in expensive hardware.¹³ The advantages of low-cost VR are numerous. First, it makes VR accessible to a broader audience, which not only means more people can experience VR but also opens up new opportunities for its use in fields such as education and training. In addition, low-cost VR can provide expanded access to education and greater flexibility for students to learn at their own pace. This can help address workforce shortages in certain professions and alleviate pressure on students.²⁰

However, low-cost VR faces challenges. Limited screen resolution and processing power can reduce immersion and experience quality.³ In addition, compared to costly options, these devices may lack advanced features like precise motion tracking and sophisticated controls, hindering natural interaction within virtual environments and limiting VR's immersive potential.³

Escape-Room Games in VR

Escape room games are interactive adventures where players work together to solve puzzles and complete challenges within a set time to "escape" from a themed room. These games often require physical interaction with objects and clues in the room. With VR, escape rooms have moved into digital spaces, allowing players to immerse themselves in virtual environments using VR headsets and controllers. This shift to VR offers more diverse and visually stunning scenarios that would be difficult to create in real life. VR escape rooms can include intricate puzzles and dynamic storytelling, enhancing the overall experience.

Several studies have explored the exciting world of VR escape rooms. Research has addressed the challenges of creating VR escape rooms, especially for mobile devices, to ensure a truly immersive experience.⁴ In addition, investigations have delved into the emotional impact of VR escape rooms, focusing on how game design can elicit specific emotions in

players.¹⁶ The potential of VR escape rooms for education has also been explored, examining their effectiveness in enhancing student motivation and critical thinking skills.¹⁷

VR escape rooms have also become a popular and diverse field within immersive gaming, offering various themes and experiences, from the intricate puzzles of “A Fisherman’s Tale”^a and “I Expect You to Die”^b to the mysterious, atmospheric challenges of “Nevrosa: Prelude.”^c Ubisoft has significantly contributed to this genre with multiplayer VR escape games such as “Escape the Lost Pyramid” and “Prince of Persia: The Dagger of Time,” set in well-known gaming worlds and optimized for entertainment venues.^d Innovative formats such as “Moonshot” incorporate AI assistants and live hosts for remote team-building, demonstrating the potential for VR to create engaging, interactive environments.^e

GPT in Immersive Technologies

Scant research has been dedicated to exploring the integration of GPT with VR and AR technologies. These recent studies have ventured into this interdisciplinary domain, uncovering the potential advantages it holds across various domains.

In a notable study by Nguyen et al.¹⁵ the focus lies in examining how the amalgamation of VR, AR, and GPT could enhance the learning experience in biology education. Their research introduced an innovative setup that allows students to interact with GPT within a simulated environment facilitated by VR/AR equipment. The findings strongly suggested that the fusion of GPT with VR/AR significantly heightened student engagement, fostering more profound and effective learning outcomes.

The study by AIContentfyTeam¹ delved into the integration of GPT within AR environments, aiming to enhance the realism of user interactions. Their research demonstrated that integrating GPT with AR could significantly elevate the authenticity and appeal of user experiences, fostering a more immersive interaction paradigm.

Moreover, Saini’s research¹⁸ shed light on the potential of GPT in VR scenarios, particularly in generating realistic dialogs within VR environments. This promising avenue showcases the capability of GPT to

enrich and augment the immersive qualities of VR experiences. In addition, in Emiliusvgs,⁵ the authors meticulously detailed the technical integration of GPT with A-frame, elucidating the methodologies employed to create VR experiences integrated with GPT via web-based platforms. Finally, a publication presented by Nandbox¹⁴ focused on delineating the enhancements offered by GPT 4 in VR experiences. This study may have highlighted the unique features of GPT 4 that render it particularly suitable for integration within VR applications.

The expanding field of LLMs is starting to intersect with immersive technologies. While research directly referencing specific LLMs in VR/AR is still developing, there is significant exploration into the potential applications of these powerful language models. One area of interest is their use as virtual tutors, with studies such as one by García-Méndez et al.⁷ highlighting their capabilities in crafting personalized learning experiences through question generation, explanation creation, and content adaptation—all crucial for effective VR/AR education. Beyond education, LLMs have the potential to revolutionize how we interact with VR/AR environments. Gatti et al.⁸ explored how LLMs can enable intuitive and natural language-based interaction within VR worlds. This could involve using voice commands to manipulate objects, navigate virtual spaces, or even engage in conversations with virtual characters. In addition, the ability of LLMs to generate creative text formats such as poems, code, scripts, and musical pieces opens doors for VR/AR applications in artistic creation and design. For instance, a concept such as that by Konenkov et al.¹¹ explores the idea of a VR system utilizing a visual language model for user interaction through NLP within VR.

Bozkir et al.² suggested using advanced LLMs in VR characters or stories. By customizing these models to users, VR could be more inclusive and engaging. However, privacy concerns exist when combining user data with VR sensor information. The article proposes LLMs in VR as a promising new research area.

In summary, these studies collectively underscore the potential transformative impact of merging GPT with VR and AR technologies across diverse domains, offering promising avenues for enhanced learning and immersive experiences. However, to the best of our knowledge, the application of GPT within the realm of low-cost VR escape rooms remains an unexplored area in the current research literature. The integration of GPT in this domain could potentially revolutionize interactive storytelling and problem-solving experiences, offering a novel dimension to immersive entertainment and

^a[Online]. Available: <https://afishermanstale-game.com/>

^b[Online]. Available: <https://iexpectyoutodie.schellgames.com/>

^c[Online]. Available: https://gexagon.com/rayden_portfolio/nevrosa-escape-en/

^d[Online]. Available: <https://www.ubisoft.com/en-us/entertainment/parks-experiences/escape-games>

^e[Online]. Available: <https://www.tryreason.com/moonshot/>

educational engagement within accessible VR environments. Further exploration and experimentation in this specific area hold promise for unlocking innovative and affordable avenues in immersive technology-driven experiences.

GPT and Logical Problems

Virtual escape room games have gained popularity as immersive entertainment experiences, challenging participants to solve puzzles and navigate through intricate scenarios. These games often require logical reasoning and problem-solving skills, prompting researchers to explore the integration of AI models, such as GPT, into these environments. Understanding the logical problem-solving capabilities of AI systems like GPT within the context of escape room challenges becomes crucial for enhancing user experiences and advancing the synergy between AI and interactive gaming environments.

Liu et al. investigated the logical problem-solving abilities of GPT, GPT-4, and RoBERTa^f across various datasets, finding GPT's superiority over RoBERTa in most benchmarks, though it struggled with newly released datasets beyond natural language inference distributions.¹² Meanwhile, Sridhara et al. explored GPT's potential in software engineering tasks, noting its remarkable performance, occasional outperformance of human experts, and conventional methods. However, they highlighted GPT's limitations and inaccuracies in certain tasks, limiting its practical applicability in those contexts.¹⁹

These investigations reveal GPT's commendable performance in logical reasoning tasks while highlighting crucial limitations, emphasizing the imperative for further advancements to broaden its efficacy across diverse problem-solving domains within virtual escape room games and beyond.

ESCAPEGPT

This section presents "EscapeGPT," a low-cost VR system that enables users to participate in solving an escape room, collaborating with GPT through voice commands. The objective is to assess whether communication between a user and an AI assistant can be truly effective in solving logical problems and challenges. The application was designed to evaluate GPT's performance in addressing highly

complex challenges and providing guidance in less demanding problem-solving situations within a specific context in the virtual environment. The developed project is freely available in a public repository.^g

Implementation

We used Unity^h for the development of EscapeGPT. Also, we used Google Cardboard libraryⁱ for low-cost VR development. Unity is a widely recognized platform used in VR application development, ensuring solid support and an active community willing to address challenges. Its seamless integration with Google Cardboard facilitates the creation of high-quality immersive experiences, aligning with our goal of developing a low-cost VR application.

Regarding the libraries used to integrate speech-to-text and text-to-speech functionalities, we selected the "Speech and Text Unity iOS and Android" library,^j which provides an efficient and straightforward solution to enable these functionalities in both Android and iOS environments. The integration process of this library is characterized by its simplicity, involving solely the importation of the corresponding packages into the development environment. Once this integration is done, access to the features provided by the library is acquired, enabling the effective implementation of voice-to-text and vice versa conversion.

Communicating with GPT required utilizing OpenAI's API,^k which stands out for its detailed and well-structured documentation, aiding in understanding how to use its capabilities.

Virtual Environment

The virtual setting was developed in Unity3D and consists of three rooms set in a manner reminiscent of a medieval dungeon, with dim lighting. These rooms can be seen in Figures 1–3. A top view of the whole virtual scenario can be seen in Figure 4.

A wide range of items can be discovered inside the rooms, including books, axes, colorful banners, wooden crates, skeletal components, tables, chairs, among others. All the assets comprising the virtual scenarios were obtained from Unity's "Asset Store."^l

^g[Online]. Available: <https://github.com/alanrychert/escapegpt>

^h[Online]. Available: <https://unity.com>

ⁱ[Online]. Available: <https://arvr.google.com/intl/cardboard/>

^j[Online]. Available: <https://github.com/j1mmyto9/Speech-And-Text-Unity-iOS-Android>, developed by Jimmy To.

^k[Online]. Available: <https://openai.com/blog/openai-api>

^l[Online]. Available: <https://assetstore.unity.com/>

^fRoBERTa: An optimized method for pretraining self-supervised NLP systems. [Online]. Available: <https://ai.meta.com/blog/roberta-an-optimized-method-for-pretraining-self-supervised-nlp-systems/>

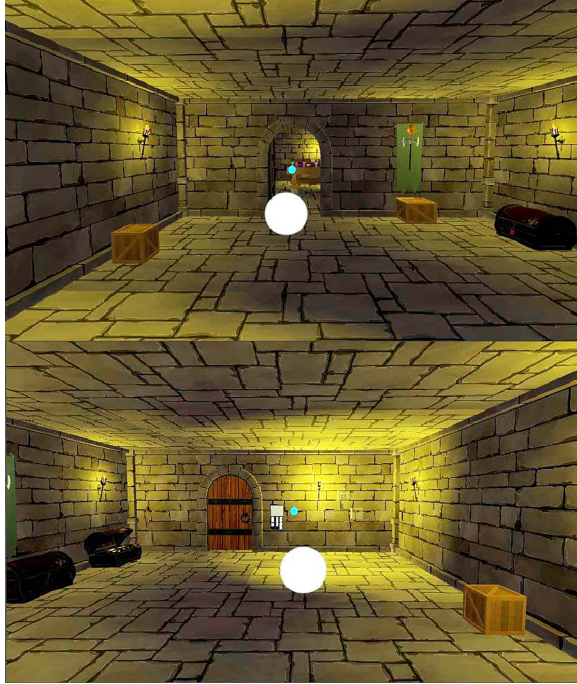


FIGURE 1. Room 1 in EscapeGPT.

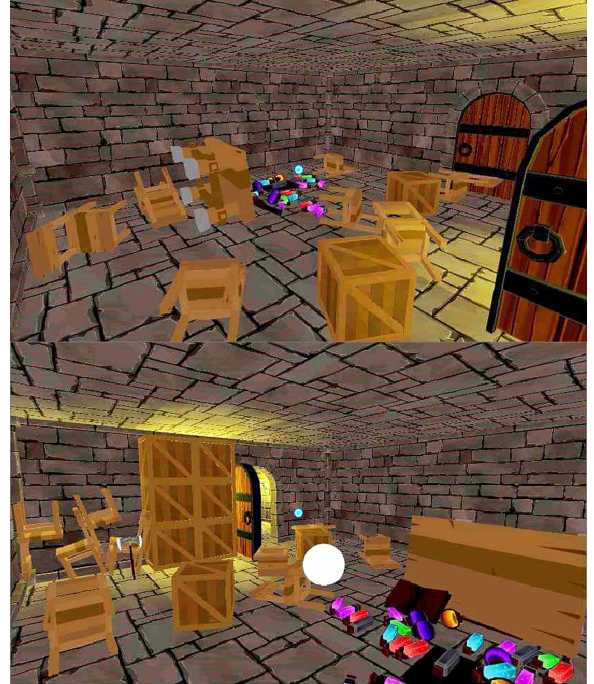


FIGURE 3. Room 3 in EscapeGPT.

Puzzles

As part of the designed escape room, several puzzles were implemented. To exemplify the average difficulty

level of these puzzles, some of them are described as follows.

In the first room, the user encounters three chests, three glass display cases, and two decorative (noninteractive) axes hanging on the wall, one



FIGURE 2. Room 2 in EscapeGPT.



FIGURE 4. Top view of the complete virtual scenario.



FIGURE 5. Main components of the first room. (a) Chests, torches, and the exit door. (b) Glass display cases. (c) Exit door with a digital padlock. (d) Axe on the wall with a skull over it. (e) Movable chest. (f) Door to the second room.

of which has a skull above it. In addition, there are multiple torches, a door, a digital padlock with a 3-digit code, along with a couple of differently sized boxes. Most of these components are shown in Figure 5.

When interacting with one of the chests marked with a red cross, it opens, revealing a book inscribed with “gleam at the door of flames.” This message refers to the torch located on the same wall as the closed door, which is the only torch in the room that can be grabbed. For instance, the user may ask the assistant about the potential interpretation of this message.

The grabbable torch has two gold coins behind it, related to the other two chests. One of the chests is open and contains gold coins, while the other is closed. When interacting with the closed chest, it pushes in and reveals a small axe that can be grabbed.

Each glass display case has a symbol, and one of them displays a key symbol and contains a lock inside since the beginning of the game. This suggests that an object related to its symbol should be placed in each showcase. Another showcases a sword symbol, indicating that the found small axe should be placed there. The third one has a symbol of a bottle. The user could query the assistant regarding the significance of

the symbols concerning the objects observed within the room, for instance.

Underneath each display case, once the corresponding object is placed, a number is shown representing the number of identical objects to the container’s contents found in the room (this detail is not obvious and requires the user to identify it themselves). This is valid for two out of the three cases, with the one supposed to contain the small axe being the exception, as instead of a number, it shows a “#” symbol to indicate that the user should count how many axes are in the room.

On the skull located above the decorative axe, there is a triangular symbol emerging from the skull’s eyes. Behind the skull, there is mirrored text that says “looks dead.” These are details emphasizing the importance of observing the position of the skeleton with red eyes found in the second room. For example, the user might enquire the assistant about this particular message or any additional hints pertaining to the skeletons.

The puzzles within the virtual escape room were meticulously crafted to strike a balance, aiming not to be overly simplistic yet avoiding extreme complexity. Their design was centered around engaging the user’s cognitive abilities, encouraging interaction with GPT for assistance. By presenting challenges that were



FIGURE 6. Final version of the assistant: A floating white sphere that follows the user.

neither excessively easy nor excessively difficult, the intention was to stimulate the user's curiosity and prompt them to seek guidance from GPT within the game environment. However, while the puzzles were strategically constructed to prompt user engagement with GPT, there remains ample room for further enhancements and refinements. Section *Discussions* addresses these aspects in more detail.

Virtual Assistant

The scenario features a virtual assistant. Initially, interaction with the assistant was done through two fixed buttons on a wall in the room. However, after conducting initial tests, it was concluded that this alternative was not intuitive for the user and did not allow for the game's continuity since the assistant's location was fixed, requiring the user to approach it for each query. Therefore, it was decided to present the assistant in a more user-friendly manner, as a sphere of light floating at a cautious distance above the ground, as shown in Figure 6. This assistant accompanies the user at all times, following the users' location but maintaining a reasonable distance from them. This design avoids the need for the user to move to specific points in the room to make inquiries.

When users interact with this assistant, they communicate through voice commands to request guidance in solving puzzles and challenges present in the escape room or to ask questions they deem useful. The assistant processes these requests and offers oral responses, adapting to the situation. Ideally, questions should include specific details, such as descriptions of objects or situations in the escape room, aiming to assist the assistant in interpreting the

context and relating key elements. This enables the assistant to utilize logical reasoning to provide useful responses and help the user overcome challenges.

Interactions

We decided to limit the available interactions to specific actions such as grabbing, dropping, pushing, opening, and making certain objects interact with each other. These interactions were considered fundamental within the context of an escape room, where challenges are based on solving logical puzzles and deducing solutions from clues present in objects and the environment. This restriction in interactions was designed to test logical reasoning and GPT's ability to understand and respond to specific queries.

The user has five different ways to interact with virtual objects:

- › **Grab:** Allows the user to pick up an object in their "virtual hand," facilitating interaction with other elements if necessary.
- › **Drop:** Enables the user to release an object by throwing it forward.
- › **Push:** This action applies force in the direction the user is looking. The intensity of this force varies depending on the object being pushed.
- › **Open:** Initiates the process of opening the interacted object.
- › **Combine:** When the user's "virtual hand" is not free, interaction with the appropriate object can combine elements and/or triggering specific events. For example, users can interact with a specific door and open it if they are currently grabbing the corresponding key.

The user has a crosshair in the center of the screen, which will be used as a pointer. When it is at a close distance and in a straight line with any interactive element in the room, text indicating the possibility of interaction becomes visible. Furthermore, the users have a joystick that allows interaction in two different ways. On one hand, they can use the stick to navigate within the rooms. On the other hand, they can use a particular button to interact with objects. The same button will be used for the five different types of interactions and for communication with the assistant. Moreover, when the user is looking at the assistant and presses the button, the sphere glows as seen in Figure 7 (top), indicating that the assistant is listening to the user's query. Users should press the button again to indicate they have finished speaking, at which point the sphere will stop glowing as seen in Figure 7 (bottom), and await the assistant's response.



FIGURE 7. Brighter sphere, indicating the assistant is listening to a query (top). Dimly lit sphere, indicating the assistant is not listening to a query (bottom).

Finally, interaction with another button was included to interrupt the assistant's response, providing the user with this option in case the response is too long and the user finds it useless.

Communication With GPT

Every time the user detects a new object in the scene, a description of that object is stored. Hence, the assistant is aware at all times of the objects the user has seen. In addition, the system maintains an initial text string about certain instructions that GPT has to know (such as the list of actions the user can perform). This text string is concatenated to the final prompt.

For example, suppose the user saw a box, a wall torch, a closed chest, and a book stating "gleam at the door of flames," and then asked the assistant, "Do you think the text 'gleam at the door of flames' could be a

clue?" Thus, the final query sent to GPT would appear as follows:

"Context: You'll act as a companion to a user trying to solve this escape room. It's essential to emphasize that the user can only grab, drop, push, open, and place objects onto others. No other actions are permitted. Not all objects are interactive. The user has only seen the following objects listed below. If the same object appears more than once, it's because the user saw more than one in the room: *box, lit wall torch, closed chest, book with the text 'gleam at the door of flames.'* Your task as a companion: assist the user in solving the escape room based on the given context but without explicitly mentioning it. **IMPORTANT:** respond to the user, not to this entire prompt. The user's inquiry is: *Do you think the text 'gleam at the door of flames' could be a clue?* Since you don't know what interactions can be done with each object, feel free to suggest sensible things, but never assure the user that they can do something with a specific object."

This message is transcribed from voice to text using the aforementioned library and sent to GPT using the corresponding API. Then, it provides us with a response, which in this example could be:

"The text 'Gleam at the door of flames' could be an important clue as it seems to stand out among the objects you've seen. It might indicate the location of a door or something related to a flame in the room. Perhaps we should further investigate the room for clues that match the text or see if there's any object that can be combined with this clue in some way. What do you think if we start by thoroughly examining the environment for additional clues?"

Once the response is received, the same library is used to transcribe this response from text to voice, thus maintaining a natural communication or dialog between the user and the assistant.

Creating an easy and smooth interaction was seen as crucial to encourage users to engage with GPT. We believed that if communicating with the system was difficult or uncomfortable, it might discourage users and make them less likely to ask questions. Therefore, our goal was to make the experience user-friendly to ensure active participation and ongoing interaction in the study.

EVALUATION

We designed an evaluation to test the usability of the system, the difficulty of the puzzles, and the effectiveness of the assistant as a companion in the room. Specifically, we are particularly interested in the capacity

TABLE 1. Results from postexperiment questionnaire—participant responses.

Question	Answers				
	Yes		No		
1. Do you play puzzle games regularly? (1 time a week or more)	12.5%		87.5%		
2. Do you have any experience in escape rooms?	37.5%		62.5%		
3. Have you ever used a virtual reality application?	75%		25%		
4. Did you manage to complete the escape room within the 25 min limit?	0%		100%		
5. Did you have problems identifying objects that could be interacted with?	50%		50%		
6. Did you need to make queries to GPT?	87.5%		12.5%		
7. Did you ever experience difficulty getting GPT to listen to you or respond to you?	25%		75%		
8. Do you think GPT adequately understood your voice queries?	100%		0%		
9. Was there a time when you felt stuck and ChatGPT helped you solve a puzzle or move forward?	75%		25%		
	Very Easy	Easy	Regular	Hard	Very Hard
10. How would you rate the difficulty of the escape room?	0%	0%	25%	50%	25%
	Yes		Sometimes		No
11. Did you find GPT’s answers helpful in moving forward in the room?	25%		62.5%		12.5%

of GPT to provide clear answers to guide the user in solving the puzzles and finding the exit of the room.

The test was conducted by eight users aged between 20 and 35 years old. It involved using the system to find the exit of the escape room while asking the assistant for any help required. Each participant had a maximum of 25 min to escape the room. Then, they had to respond to a questionnaire that included questions related to their prior experience using VR, how frequently they play puzzle games, their perception of the system’s usability, their opinion about GPT as a companion in the room regarding the quality and usefulness of its responses, and any other information users considered relevant. The results from the postexperiment questionnaire and participant responses are presented in Table 1.

Procedure

Before starting the virtual experience, the participants were asked to read a small guide that contained a brief explanation about the escape room, the purpose of the test, and instructions detailing the equipment to be used, controls, and methods of interaction so that the user could navigate smoothly within the virtual environment, whether interacting with objects, moving around, or communicating with the assistant represented by the floating sphere.

Participants had time to put on and adjust the VR headset, which they did not remove until the end of the experiment. Throughout the experiment, they were accompanied by a researcher to provide any necessary help or physical support.

Throughout the test, the responsible researcher took notes on the participant’s behavior, noting, among other things, each user’s achievements and the general format of the questions the participant asked GPT. After 25 min or once the participants completed the escape room, they were asked to complete the mentioned questionnaire.

Results

A total of eight users participated in the evaluation of the VR escape room with the assistance of GPT, comprising five men and three women. Among these, one user claimed to regularly play puzzle games, while three had previous experience with escape rooms. In addition, six participants had used VR systems on at least one previous occasion.

Regarding the difficulty of the escape room, half of the participants considered it “hard,” 2 rated it as “very hard,” and 2 as “regular.” It is noteworthy that none of the users managed to escape from the room, highlighting the high difficulty of the experience. Also, during the game, half of the participants had difficulty identifying the interactive objects in

the escape room, suggesting a potential improvement in visual feedback.

All the participants needed to make questions to GPT for guidance or puzzle-solving, and all considered that the assistant adequately understood their queries. However, only 2 of the participants experienced difficulties with GPT responding or hearing their questions, which was due to an internet connection issue that hindered smooth communication with the assistant. Regarding the perception of GPT, five participants stated that the assistant's responses were sometimes useful, two participants felt that their responses were always useful, while one indicated that the responses received were never useful.

As mentioned, the responsible researcher took detailed notes while the participants were undergoing the evaluation. The researcher paid special attention to the participants' behavior and the questions they posed to the GPT assistant. The experiment reveals several important aspects regarding usability, human-computer interaction, and the current limitations of technology in VR environments.

At the beginning of the experiment, several users expressed confusion about how to interact with the GPT assistant. For example, one user asked, "What's going on?" and "I need hints." This finding underscores the need to improve initial instructions and the user interface to facilitate a quicker and more efficient understanding of the system. Clarity in instructions is crucial to reduce the learning curve and avoid early frustrations that can dissuade users.

Throughout the experiment, participants asked a variety of questions to the GPT assistant, highlighting different levels of understanding and strategies for solving the presented problems. For instance, some users asked about the use of specific objects, such as "What color of potion is the most useful?" "What do I do with the potions?" and "Where did I leave the axe?" Other users tended to ask very specific questions about objects and their possible interactions, such as "How do I grab the key to the door?" "What do I do with the book I just grabbed?" and "Would the chest with coins help solve something in the escape room?"

The experiment also revealed significant technical limitations. For example, one user experienced problems with Wi-Fi quality, which affected the screen-sharing function and thus the overall user experience. In addition, the text-to-speech library used showed limitations in question intonation, which could affect the perception of naturalness and effectiveness of the assistant.

The performance of the GPT assistant varied among users. Although in some cases it provided

useful answers that helped users progress, in other cases the responses were perceived as vague or insufficient. For example, when a user asked, "What can I do to get out of the room?" the response was hindered by connection issues. However, in other instances, the GPT assistant provided helpful answers such as suggesting, "You might need to find a key to unlock the door," or guiding users on how to interact with specific objects, like when asked, "How do I get the key out of there?" it responded, "Try using a tool to reach the key." For questions like "What does gleam at the door of flames have to do with it?" the assistant clarified, "Fulgur likely refers to something glowing, such as a torch." And when asked, "What can I try to burn with the torch?" GPT suggested, "You could try burning objects that seem flammable, like a web or a wooden barrier." This suggests that, while GPT is a powerful tool and often provided useful guidance, it still has limitations in its ability to understand and respond effectively in specific VR contexts.

Overall, some participants experienced issues with in-game controls and dizziness. Despite the challenges, the experience left participants engaged and eager to continue, highlighting the need for wit and patience to solve the puzzles in the escape room.

DISCUSSION

We conducted a user study to evaluate participants' perceptions of a virtual assistant integrated with GPT, while also assessing the user experience within a low-cost VR setup. This assistant was designed to assist users in navigating an escape room containing multiple puzzles. Users could interact with the assistant by providing descriptions of objects, potential actions, and receiving a brief overview of the escape room and their objectives. Our objective was to assess the effectiveness of the assistant's responses, users' progress in solving the puzzles, and the overall user experience within the low-cost VR environment. This section presents a discussion of these findings, acknowledges the limitations of our study, and proposes avenues for future research.

The lack of success in solving the escape room by any of the users could be attributed to several factors. First, the predominant rating of the room as "very difficult" suggests that the challenges presented are considerably complex, requiring a high degree of wit and logical reasoning. This indicates that the escape room posed a significant challenge not easily overcome.

In addition, participants' lack of previous experience in escape rooms and limited familiarity with VR applications could have contributed to the difficulty in

solving the escape room. Lack of knowledge and practice in these specific environments might have influenced their ability to tackle the challenges presented. However, participants provided constructive suggestions to enhance the interactions with GPT, including the idea of providing shorter responses and the assistant having more information about the room.

Since none of the participants managed to solve the escape room, there is room for adjusting the difficulty of puzzles and riddles. In addition, the inclusion of multiple difficulty levels in the escape room could appeal to a broader audience, ranging from beginners to puzzle game enthusiasts and escape room experts.

In summary, the usability test provided valuable qualitative user feedback. In future iterations, collecting more detailed quantitative data, such as the time users take to solve puzzles or the frequency of queries to GPT, would be beneficial to evaluate the performance and experience of both users and the assistant.

Recommendations for future improvements include enhancing the user interface through interactive tutorials and clearer initial instructions, optimizing the quality of internet service and text-to-speech technology integration, and considering users' physical well-being by incorporating scheduled alerts and breaks. Additionally, it is essential to improve the GPT assistant's ability to provide more specific and contextually relevant responses, possibly through additional training in VR scenarios.

The study demonstrates that while the integration of a GPT assistant in a virtual reality environment has great potential, there are still significant areas that require improvement to optimize the user experience. The findings provide a solid foundation for future research and developments in this emerging area of human-computer interaction.

It is important to note limitations inherent in using the free trial version of GPT API. First, the API lacks conversation context maintenance between requests, treating each query as an independent unit without knowledge of previous interactions. Thus, the client application must provide complete context in each request to ensure conversation coherence. In addition, access to the latest model updates is restricted, potentially affecting its understanding of certain contexts or providing updated responses. Moreover, the free version imposes token limits per query, addressed by a predefined prompt format including essential conversation information and user-provided details. Considering these limitations is crucial when implementing the GPT API, as they may impact interaction and results obtained. Future iterations could benefit

from using an advanced version to leverage improvements in handling longer, more complex conversations. Utilizing a paid version like GPT-4 could also prove advantageous.

CONCLUSION

In conclusion, the integration of GPT as an assistant in low-cost virtual escape games shows promise in enhancing problem-solving experiences within the virtual environment. User evaluations revealed both positive perceptions and limitations, highlighting the potential of GPT as a valuable tool for providing assistance and guidance in complex challenges. Areas for future improvement include adjusting puzzle difficulty and enhancing GPT's contextual understanding. This research sheds light on the opportunities and challenges of incorporating AI language models like GPT in virtual gaming environments, paving the way for further advancements in this field.

ACKNOWLEDGMENTS

This work was supported in part by the Secretaría General de Ciencia y Tecnología, Universidad Nacional del Sur (Argentina) under Grant PGI 24/ZN38, Grant PGI 24/ZN42, and Grant PGI 24/ZN43 and in part by the National Council for Scientific and Technical Research under Grant 28720210100824CO (PIBAA).

REFERENCES

1. AIContentfyTeam, "Using ChatGPT for creating engaging and interactive augmented reality experiences," Accessed: Aug. 2, 2024. [Online]. Available: <https://aicontentfy.com/en/blog/chatgpt-for-creating-engaging-and-interactive-augmented-reality-experiences>
2. E. Bozkir, S. Özdel, K. H. C. Lau, M. Wang, H. Gao, and E. Kasneci, "Embedding large language models into extended reality: Opportunities and challenges for inclusion, engagement, and privacy," in *Proc. 6th ACM Conf. Conversational User Interfaces*, 2024, pp. 1–7.
3. J. Q. Coburn, I. Freeman, and J. L. Salmon, "A review of the capabilities of current low-cost virtual reality technology and its potential to enhance the design process," *ASME. J. Comput. Inf. Sci. Eng.*, vol. 17, no. 3, Jul. 2017, Art. no. 031013, doi: [10.1115/1.4036921](https://doi.org/10.1115/1.4036921).
4. C. C. De Araújo, J. F. H. de Melo, J. M. X. N. Teixeira, and F. F. F. Peres, "Developing a virtual escape room game: Existing challenges and how to overcome them," in *Proc. Anais Do XX Congresso Latino-Americano De Softw. Livre E Tecnologias Abertas*, Oct. 2023, pp. 40–49, doi: [10.5753/latinoware.2023.236402](https://doi.org/10.5753/latinoware.2023.236402).

5. Emiliusvgs, "ChatGPT—Open IA creates web augmented reality with a-frame," Accessed: Aug. 2, 2024. [Online]. Available: https://www.youtube.com/watch?v=4Y1yzlAKBvY&ab_channel=Emiliusvgs
6. X. Ding and Z. Li, "A review of the application of virtual reality technology in higher education based on Web of Science literature data as an example," *Front. Educ.*, vol. 7, 2022, Art. no. 1048816, doi: [10.3389/feduc.2022.1048816](https://doi.org/10.3389/feduc.2022.1048816).
7. S. García-Méndez, F. de Arriba-Pérez, and M. D. C. Somoza-López, "A review on the use of large language models as virtual tutors," *Sci. Educ.*, 1–16, 2024, doi: [10.1007/s11191-024-00530-2](https://doi.org/10.1007/s11191-024-00530-2).
8. A. Gatti, "Virtual environments via natural language agents," in *Proc. Eur. Conf. Multi-Agent Syst.*, Cham: Springer Nature Switzerland, Sep. 2023, pp. 486–492, doi: [10.1007/978-3-031-43264-4_36](https://doi.org/10.1007/978-3-031-43264-4_36).
9. A. Ioannou, M. Lemonari, F. Liarokapis, and A. Aristidou, "Collaborative VR: Solving riddles in the concept of escape rooms," in *Proc. Int. Conf. Interactive Media, Smart Syst. Emerg. Technol.*, 2023, pp. 1–9, doi: [10.2312/imet.20231247](https://doi.org/10.2312/imet.20231247).
10. N. Jagdishbhai and K. Y. Thakkar, "Exploring the capabilities and limitations of GPT and chat GPT in natural language processing," *J. Manage. Res. Anal.*, vol. 10 no. 1, 18–20, 2023, doi: [10.18231/jjmra.2023.004](https://doi.org/10.18231/jjmra.2023.004).
11. M. Konenkov, A. Lykov, D. Trinitatova, and D. Tsetserukou, "VR-GPT: Visual language model for intelligent virtual reality applications," 2024, *arXiv:2405.11537*.
12. H. Liu, R. Ning, Z. Teng, J. Liu, Q. Zhou, and Y. Zhang, "Evaluating the logical reasoning ability of ChatGPT and GPT-4," 2023, *arXiv:2304.03439*.
13. H. E. Lowood, "Virtual reality," *Encyclopedia Britannica*, Nov. 2023. [Online]. Available: <https://www.britannica.com/technology/virtual-reality>
14. Nandbox. "ChatGPT 4 and virtual reality: Enhancing immersive experiences," Accessed: Aug. 2, 2024. [Online]. Available: <https://nandbox.com/chat-gpt-4-and-virtual-reality-enhancing-immersive-experiences/>
15. T. Nguyen, L. Cao, and H. Tru'ong, "Integrating ChatGPT with virtual reality and augmented reality technology in teaching biology at high schools in Vietnam," 2022, doi: [10.35542/osf.io/aezvw](https://doi.org/10.35542/osf.io/aezvw).
16. I. Oliveira, V. Carvalho, F. Soares, P. Novais, E. Oliveira, and L. Gomes, "Development of a virtual reality escape room game for emotion elicitation," *Information*, vol. 14, no. 9, 2023, Art. no. 514, doi: [10.3390/info14090514](https://doi.org/10.3390/info14090514).
17. J. M. P. Piernas, M. C. P. Meroño, and M. D. P. F. Asenjo, "Virtual escape rooms: A gamification tool to enhance motivation in distance education," *RIED-Revista Iberoamericana de Educación a Distancia*, vol. 27, no. 1, pp. 61–85, 2024, doi: [10.5944/ried.27.1.37685](https://doi.org/10.5944/ried.27.1.37685).
18. N. Saini, "The use of ChatGPT in virtual reality experiences," Aug. 2, 2024. [Online]. Available: <https://www.linkedin.com/pulse/use-chatgpt-virtual-reality-experiences-dr-nitin-saini/>
19. S. Sridhara, S. Kambhampati, and H. Liu, "ChatGPT: A study on its utility for ubiquitous software engineering tasks," 2023, *arXiv:2305.16837*.
20. K. Walsh et al., "High value and low-cost virtual reality healthcare professional education: Proceedings of a roundtable workshop," *MedEdPublish*, vol. 9, pp. 9–57, 2020, doi: [10.15694/mep.2020.000057.1](https://doi.org/10.15694/mep.2020.000057.1).

ALAN RYCHERT is an information systems engineer and teaching assistant with the Department of Computer Science and Engineering, Universidad Nacional del Sur, Bahía Blanca, 8000, Argentina. His research interest focuses on the study of different interactions of low-cost virtual and augmented reality. Contact him at alanrychert@gmail.com.

MARÍA LUJÁN GANUZA is a researcher with the Institute for Computer Science and Engineering and a professor with Institute for Computer Science and Engineering (UNS-CONICET), and VyGLab Research Laboratory (UNS-CICPBA), Department of Computer Science and Engineering, Universidad Nacional del Sur, Bahía Blanca, 8000, Argentina. She is currently working on the development of visualization techniques for multidimensional, time-oriented, and spatial data, focusing on the development of visualization techniques for geological data and eye-tracking data. Contact her at mlg@cs.uns.edu.ar.

MATIAS NICOLÁS SELZER is a researcher with the Institute for Computer Science and Engineering and a professor with the Institute for Computer Science and Engineering (UNS-CONICET), and VyGLab Research Laboratory (UNS-CICPBA), Department of Computer Science and Engineering, Universidad Nacional del Sur, Bahía Blanca, 8000, Argentina. He is a member of the VyGLab Research Laboratory since 2015. His research interest focuses on the study of presence and immersion in virtual reality, augmented reality, and the relation between virtual reality and the human senses. He is the corresponding author of this article. Contact him at matias.selzer@cs.uns.edu.ar.