Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 20 No. 15 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v20i15.51437

PAPER

In-World NPC: Analysing Artificial Intelligence Precision in Virtual Reality Settings

Mochamad Bruri Triyono[1], Arif Ainur Rafiq[2](✉), Didik Hariyanto[3], Dina Adinda[4], Maria Denami[5]

[1]Department of TVET, Graduate School, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

[2]Department of Mechatronics Engineering, Politeknik Negeri Cilacap, Cilacap, Indonesia

[3]Department of Electrical Engineering Education, Faculty of Engineering, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

[4]CREF – UR 1589, Université Paris Nanterre, France

[5]LISEC – UR 2310, University of Upper Alsace, France

aar@pnc.ac.id

ABSTRACT

This study investigates the accuracy and performance of artificial intelligence (AI)-driven non-player characters (NPC) in a virtual classroom environment, focusing on an educational simulation scenario. By employing a mixed-methods approach that combines quantitative data from controlled experiments and qualitative insights from user feedback, the research aims to provide a comprehensive evaluation of NPC behavior. The findings indicate that NPC demonstrated a high decision-making accuracy rate of 87%, with solid behavioral consistency and generally fast response times averaging 1.2 seconds. However, challenges were noted in handling complex queries and maintaining consistency in dynamic scenarios. Factors influencing NPC performance included the complexity of user interactions, the environmental context, and the limitations of AI algorithms. Overall user satisfaction was high, with participants appreciating NPC realism, responsiveness, and engagement. A comparative analysis of AI techniques revealed that while rule-based systems were efficient and predictable, machine learning models offered superior adaptability and contextual understanding, albeit at the cost of higher computational demands. The study concludes that enhancing AI capabilities, optimising computational resources, and incorporating adaptive learning algorithms can improve NPC performance. These insights provide valuable guidance for future developments in AI-driven NPC, aiming to create more immersive and compelling virtual educational environments.

KEYWORDS

artificial intelligence (AI) accuracy, non-player characters (NPC), virtual reality (VR), NPC behavior, virtual reality (VR) scenarios

1 INTRODUCTION

Virtual reality (VR) technology has rapidly evolved over the past few years, transforming from a niche entertainment medium into a versatile tool across various domains, including education, training, healthcare, and social interactions [1], [2], [3]. This evolution has been driven by significant advancements in hardware and software, enabling more immersive and interactive user experiences. A crucial

Triyono, M.B., Rafiq, A.A., Hariyanto, D., Adinda, D., Denami, M. (2024). In-World NPC: Analysing Artificial Intelligence Precision in Virtual Reality Settings. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(15), pp. 19–42. https://doi.org/10.3991/ijoe.v20i15.51437

Article submitted 2024-07-29. Revision uploaded 2024-09-24. Final acceptance 2024-09-27.

© 2024 by the authors of this article. Published under CC-BY.

element in creating these immersive environments is the presence of NPC, which artificial intelligence (AI) controls to interact with users in realistic ways [4].

Non-player characters (NPC) are the backbone of interactive VR environments, providing dynamic interactions that enhance the realism and depth of the virtual world [5], [6], [7]. NPC can take on various roles, from simple bystanders to complex agents that engage users meaningfully, contributing to the overall narrative and user experience. The effectiveness of these NPC heavily depends on the accuracy and sophistication of the underlying AI algorithms that govern their behavior [8].

Artificial intelligence has made significant strides in recent years, particularly in fields such as machine learning and natural language processing (NLP), which are critical for developing intelligent NPC [9]. These advancements have enabled NPC to exhibit more lifelike behaviors, such as adaptive learning, emotional responses, and context-aware interactions, thereby enhancing the immersive quality of VR experiences [10]. However, despite these advancements, challenges remain in ensuring the accuracy and reliability of NPC behavior across different VR scenarios.

The accuracy of AI in controlling NPC is vital for maintaining the illusion of reality in VR environments. High accuracy ensures that NPC behaves in ways consistent with user expectations and the virtual world's rules, preventing disruptions to the immersive experience [11], [12]. Inaccurate or erratic NPC behavior can break the sense of immersion and reduce the overall effectiveness of VR applications, whether for entertainment, training, or therapeutic purposes [13].

Previous studies have highlighted the importance of AI accuracy in VR settings, emphasizing the need for robust and adaptable AI systems that can handle the complexities of real-time interactions [6], [14], [15]. These studies have explored various AI techniques to improve NPC performance and user satisfaction, including rule-based systems, machine-learning models, and hybrid approaches [16]. However, there is still a lack of comprehensive research that systematically evaluates AI accuracy across different VR scenarios and identifies the key factors influencing NPC performance.

This study aims to fill this gap by thoroughly analyzing AI precision in managing NPC behavior within diverse VR settings. By examining multiple VR scenarios, we strive to identify the strengths and weaknesses of current AI implementations and provide insights into the factors that affect NPC accuracy [7], [17], [18], [19]. This study is essential for advancing AI technologies in VR and ensuring NPC can deliver consistent and realistic interactions across various applications.

Our research methodology involves a mixed-methods approach, combining quantitative and qualitative analyses to evaluate AI accuracy. We selected a range of VR scenarios representing different levels of interaction complexity and NPC roles, ensuring a comprehensive assessment of AI performance [20]. Data were collected through user interactions with NPC, system logs, and feedback surveys, allowing us to gain a holistic understanding of AI accuracy and user experiences.

The results of our study provide valuable insights into the current state of AI in VR and highlight areas where improvements are needed. By identifying the factors that impact NPC performance, we can develop more effective AI strategies that enhance the realism and immersion of VR environments [21], [22], [23], [24]. These findings can also inform future research and development efforts, guiding the creation of more sophisticated and reliable AI systems for VR applications.

In the following sections, we will review the existing literature on AI in VR and NPC behavior, describe our research methodology, present the results of our analysis, and discuss the implications of our findings. We also provide recommendations for enhancing AI accuracy in VR, drawing on our study to suggest practical strategies for developers and researchers [23], [25].

By advancing our understanding of AI accuracy in VR, this study contributes to the broader goal of creating more immersive and interactive virtual experiences. Whether used for gaming, education, training, or therapeutic purposes, accurate and realistic NPCs are essential for maximizing the potential of VR technology [26], [27], [28], [29]. Through this study, we aim to support the ongoing evolution of VR and help realize its full potential as a transformative medium.

2 LITERATURE REVIEW

In this literature review, we explored the current state of AI in VR, the role of NPC in creating immersive and interactive virtual environments, and the methodologies for measuring AI accuracy and performance. This examination will comprehensively understand how AI-driven NPC can be optimized to support vocational education and other applications.

2.1 AI in virtual reality

Integrating AI into VR has significantly transformed how virtual environments are created and experienced. AI technologies enable the creation of intelligent, responsive, and adaptive elements within VR, enhancing the realism and interactivity of these environments [30]. One of the primary applications of AI in VR is the development of NPC, which can interact with users in complex and meaningful ways. These NPCs are essential for creating engaging and immersive experiences in gaming, training, or therapeutic contexts [31], [32]. AI technologies used in VR range from simple rule-based systems to advanced machine learning algorithms. Rule-based systems operate on predefined sets of rules and logic to dictate NPC behavior. While these systems are relatively straightforward to implement and understand, they lack the flexibility and adaptability required for more complex interactions [33]. In contrast, machine learning algorithms, particularly those based on deep learning, can learn from data and adapt their behavior over time, making them more suitable for dynamic and unpredictable environments [34], [35], [36], [37].

Machine learning techniques, such as reinforcement learning, have been widely adopted in VR to improve NPC behavior. Reinforcement learning allows NPCs to learn optimal behaviors through trial and error, receiving feedback from the environment to refine their actions [38]. This approach is efficient in VR scenarios where NPCs must navigate complex environments, interact with users, and make decisions based on real-time data. However, the computational demands of reinforcement learning can be substantial, requiring significant processing power and time [38], [39], [40].

Natural language processing is another critical area of AI that enhances VR experiences. NLP enables NPC to understand and generate human language, allowing more natural and intuitive interactions between users and virtual characters [41]. Advances in NLP, such as transformer models, have significantly improved the accuracy and fluency of AI-driven dialogue systems, making NPC more believable and engaging [42]. These improvements have broad applications, from educational VR environments to therapeutic interventions where NPCs act as virtual therapists.

The integration of AI in VR also extends to procedural content generation, where AI algorithms create dynamic and personalized content for users. This approach allows for the creation of vast and varied virtual worlds without requiring extensive manual design [43], [44], [45]. Procedural content generation can create unique levels, environments, and scenarios, ensuring users have a distinct and tailored experience. AI-driven content generation is instrumental in gaming, where replayability and diversity are essential.

Despite the significant advancements in AI technologies for VR, challenges remain in achieving high levels of accuracy and reliability. One of the primary challenges is the computational complexity of AI algorithms, which can limit their scalability and real-time performance [34]. Additionally, the unpredictability of user behavior in VR environments poses a significant challenge for AI systems, requiring them to adapt quickly and effectively to various scenarios [46]. Ensuring that AI-driven NPC can respond appropriately to diverse and unforeseen interactions is a critical area of ongoing research.

Another challenge is the ethical considerations surrounding AI in VR. Using AI to create lifelike and emotionally engaging NPC raises questions about user privacy, data security, and the potential for manipulation [47]. Ensuring that AI systems are designed and used responsibly is essential for maintaining user trust and safeguarding against misuse. It includes implementing transparent AI practices, protecting user data, and establishing ethical guidelines for AI development and deployment in virtual reality.

Furthermore, there is a need for standardized metrics and evaluation frameworks to assess the performance and accuracy of AI in VR. Current evaluation methods often vary widely, making it difficult to compare results across studies and applications [48], [49], [50]. Developing comprehensive and consistent evaluation criteria is crucial for advancing the field and ensuring that AI systems meet the required performance and reliability standards.

2.2 Role of NPCs in VR environments

Non-player characters are crucial for creating immersive VR environments, acting as interactive agents that enhance user engagement through context, narrative, and interaction [46]. NPC can perform a wide range of functions, from simple background characters that enhance the realism of the environment to complex interactive agents that drive the story and respond to user actions [51]. By simulating human-like behavior, NPCs enhance realism, enabling users to immerse themselves fully in VR experiences [52], [53]. For instance, NPCs can act as instructors or peers in training simulations, providing feedback and supporting educational outcomes [54].

The interactivity of NPCs is critical to responsive environments, as they adapt in real time to user inputs, keeping users engaged [55]. In gaming, an NPC's reaction to player actions can create more challenging and immersive experiences [56]. NPCs drive the storyline in narrative-driven VR environments, guiding users through missions and tasks while contributing to a coherent, compelling narrative [57].

In social VR applications, NPCs facilitate interaction and provide companionship, particularly useful in therapeutic and educational contexts [25], [58]. Developing these believable NPCs requires advanced AI, including NLP, emotion recognition, and adaptive machine-learning techniques [41], [1], [34]. Despite advancements, designing flexible, robust AI systems that handle diverse user interactions remains challenging [18], [46], and computational efficiency is crucial [34]. Ethically, creating life-like NPCs raises concerns about manipulation and privacy. Responsible AI development with clear ethical guidelines, transparency, and data protection measures is essential to maintaining user trust [47], [61], [62].

2.3 Measuring AI accuracy and performance

Measuring the accuracy and performance of AI in controlling NPC within VR environments is crucial for ensuring the realism and effectiveness of these virtual worlds. Accurate AI-driven NPC enhances the immersive experience by behaving in ways consistent with user expectations and the virtual world's rules [50]. The assessment of AI accuracy involves evaluating various aspects of NPC behavior, including responsiveness, adaptability, and believability [33]. Behavioral analysis is one of the primary methods for measuring AI accuracy in NPC. It involves observing and recording NPC actions responding to user interactions and environmental changes. Behavioral metrics such as reaction time, decision-making accuracy, and adherence to predefined rules are commonly used to assess the performance of AI algorithms [55]. These metrics provide quantitative data that can be analyzed to determine the effectiveness of AI in simulating realistic behaviors.

User feedback is another essential component of evaluating AI accuracy in VR. Collecting subjective assessments from users provides valuable insights into how believable and engaging they find in NPC interactions. Surveys, interviews, and usability studies are commonly used to gather this feedback, which can then be correlated with objective behavioral metrics to comprehensively understand AI performance [13]. User feedback helps identify areas where AI-driven NPCs excel and where improvements are needed. Simulation-based testing is also widely used to measure AI accuracy in VR. It involves creating controlled scenarios where NPC are subjected to specific tasks and challenges that test their abilities. By analyzing NPC performances in these simulations, researchers can identify strengths and weaknesses in the AI algorithms [63]. Simulation-based testing allows for systematic and repeatable evaluations, making it a valuable tool for AI development and refinement.

Machine learning techniques such as reinforcement learning often include built-in mechanisms for evaluating AI performance. These techniques use reward functions to provide feedback to the AI system based on its actions, enabling continuous learning and improvement [38]. The performance of reinforcement learning algorithms can be assessed by tracking the cumulative rewards achieved over time, which indicates how well the AI adapts to the environment and user interactions. This approach provides a dynamic and ongoing assessment of AI accuracy. Another method for measuring AI accuracy is using benchmarking datasets and tasks. These standardized datasets and functions provide a common framework for evaluating and comparing AI algorithms [34]. By testing AI-driven NPC on these benchmarks, researchers can objectively assess their performance and identify best practices and areas for improvement. Benchmarking is particularly useful for advancing the state of the art in AI and ensuring that new algorithms meet established performance standards.

The evaluation of AI accuracy also involves assessing the computational efficiency of the algorithms. High-performance AI algorithms must balance accuracy with computational demands, ensuring they can operate in real-time within VR environments [64], [65]. Metrics such as processing time, memory usage, and scalability are important considerations for evaluating the practical feasibility of AI-driven NPC. Efficient algorithms are essential for maintaining a smooth and responsive VR experience. Ethical considerations are increasingly important in the evaluation of AI accuracy. Ensuring that AI-driven NPC behaves in ways that are ethical and respectful of user autonomy is critical for maintaining user trust and safeguarding against potential misuse [47]. Evaluation frameworks should include ethical criteria, such as fairness, transparency, and accountability, to ensure that AI systems are designed and used responsibly. It includes avoiding biases in AI behavior and protecting user data.

The development of standardized evaluation frameworks is essential for advancing the field of AI in VR [59], [60]. Currently, evaluation methods vary widely, making comparing results across studies and applications difficult. Standardized frameworks would provide a standard set of criteria and metrics for assessing AI accuracy, facilitating the comparison of different algorithms, and promoting best practices. This standardization is crucial for ensuring the reliability and robustness of AI-driven NPC. In addition to quantitative metrics, qualitative analysis is essential in evaluating AI accuracy. Qualitative methods, such as thematic and content analysis, provide deeper insights into user experiences and perceptions of NPC behavior [66]. These methods can uncover nuances in NPC interactions that quantitative metrics may miss, offering a more holistic understanding of AI performance and combining qualitative and quantitative approaches to evaluate AI accuracy in virtual reality.

In conclusion, measuring AI accuracy and performance in VR involves a multifaceted approach that includes behavioral analysis, user feedback, simulation-based testing, machine learning metrics, benchmarking, computational efficiency, ethical considerations, and standardized frameworks. A comprehensive evaluation framework that integrates these methods is essential for advancing the field and ensuring that AI-driven NPCs deliver realistic, engaging, and ethical interactions in VR environments. Continued research and development in these areas are critical for improving the accuracy and effectiveness of AI in virtual reality.

3 METHOD

3.1 Research design

The study adopts a mixed-methods approach to evaluate the accuracy and performance of AI-driven NPC within various VR scenarios. This approach combines quantitative data from controlled experiments with qualitative insights from user feedback, ensuring a comprehensive assessment of AI performance.

The increasing interest and potential of VR technology in educational settings drove the focus on the educational simulation scenario. Educational simulations offer a unique opportunity to explore how VR can enhance learning experiences, improve engagement, and facilitate interactive teaching methods. By concentrating on this specific scenario, the study aimed to generate insights that are directly applicable to the field of education, where VR is seen as a promising tool for innovation. Additionally, the educational simulation provided a controlled environment to systematically observe and measure user interactions and outcomes, ensuring the reliability and validity of the findings.

3.2 VR scenarios and NPC selection

One distinct VR scenario was selected for the study to represent varying levels of interaction complexity: an educational simulation featuring a virtual classroom where NPCs serve as instructors and fellow students. NPC in this scenario was designed with specific roles and behaviors relevant to the academic context, aiming to cover a broad spectrum of interaction types, from structured tasks in the classroom setting to dynamic, unpredictable interactions that might be encountered in a gaming environment.

3.3 Data collection and analysis

In this study, data collection involved two main phases: qualitative and quantitative. Table 1 describes it. This study used thematic analysis to analyze the qualitative data collected from participant interactions and feedback. The analysis was managed manually without the use of specialized software. Researchers carefully reviewed and coded the data to identify recurring themes and patterns. This manual approach allowed for a more nuanced understanding of the participants' experiences and perceptions, ensuring that subtle insights were not overlooked. The themes were then categorized and analyzed to draw meaningful conclusions about the effectiveness and user experience of the educational simulation scenario in the VR environment.

Two independent evaluators were involved in the coding process to ensure the reliability and validity of the thematic analysis. Both evaluators had extensive experience in qualitative research methods and were familiar with the subject matter of VR technology in educational settings. They independently reviewed the data, generated initial codes, and then compared and discussed to resolve discrepancies. This collaborative approach helped to enhance the credibility of the findings by ensuring that the identified themes accurately reflected the participants' experiences and perceptions. The themes were then categorized and analyzed to draw meaningful conclusions about the effectiveness and user experience of the educational simulation scenario in the VR environment.

Table 1. Data collection descriptions

Phase 1: Quantitative Data Collection	Phase 2: Qualitative Data Collection
• Controlled Experiments. Participants engaged with NPC in each VR scenario while their interactions were recorded. Key metrics included response time, decision-making accuracy, and behavioral consistency.	• User Feedback. After interacting with the NPC, participants completed surveys and participated in interviews to provide feedback on their experiences [67]. The surveys included Likert scale questions on NPC realism, responsiveness, and overall satisfaction.
• Performance Metrics. Data were analysed using statistical methods such as descriptive statistics to identify patterns and significant differences in NPC performance across scenarios.	• Thematic Analysis. Qualitative data from interviews were analysed to identify common themes and insights into user perceptions of NPC behavior [68].

3.4 Participants

The study recruited 30 participants with varying levels of familiarity with VR technology, including a balanced mix of genders, various ethnic backgrounds, and a range of educational levels from high school graduates to postgraduates. Participants were divided into three groups using random assignment, ensuring that each group had a similar distribution of ages, genders, and VR experience levels. The ages of the participants ranged from 18 to 20 years old. Participants' expertise in VR varied from beginners with no prior experience to advanced users who regularly engage with VR technology. This diverse participant pool helped ensure the findings were generalizable across user experiences and expectations.

4 RESULTS

This section presents the findings from the study focused on the educational simulation scenario, where NPC served as instructors and fellow students in a virtual classroom environment. The results are divided into two main parts: quantitative data from controlled experiments and qualitative insights from user feedback. The educational simulation scenario observed in this study was designed to replicate a virtual classroom environment where participants interacted with NPC acting as students. The scenario included various academic tasks such as delivering a lecture, facilitating a group discussion, and managing classroom dynamics. The VR environment was equipped with interactive elements like virtual whiteboards and presentation tools to enhance the realism of the educational experience. This setup aimed to comprehensively understand how users engage with VR in academic contexts, focusing on user immersion, interaction quality, and learning outcomes.

4.1 AI accuracy in the educational simulation scenario

Artificial intelligence accuracy ensures relevant and reliable outcomes in the educational simulation scenario. We have divided this explanation into three main sub-sections: response time, decision-making accuracy, and behavioral consistency. Response time refers to how quickly the AI can provide answers or actions in a given situation, essential for creating a dynamic and interactive learning environment. Decision-making accuracy relates to the extent to which the AI can make correct decisions based on the provided context. Lastly, behavioral consistency refers to the AI's ability to maintain actions that align with pre-programmed behavioral patterns, ensuring a stable and predictable learning experience.

The response time of AI-driven NPC in the educational simulation scenario was measured by recording the time for NPC to react to user inputs. This metric is crucial for ensuring that interactions between users and NPC are smooth and responsive, contributing to the overall realism and immersion of the virtual classroom environment. The data is collected in three ways:

a) Two hundred interactions between participants and NPC were recorded in the virtual classroom.
b) Each interaction involved a user input, such as a question or request, followed by an NPC response.
c) The user input time and the subsequent NPC response were precisely logged using custom scripts within the VR environment.

In Unity 3D, we created a script to log when user input is detected and when an NPC responds. The next stage is conducting the VR session with participants and logging the response for each interaction. The logging script is shown in Figure 1. After completing the session, collect the log file with all the recorded response times. To calculate the average response time, we imported the data into a Python programming environment (see Figure 2). In this paper, we assumed the following logged data in *'response_time.csv'* (see Figure 3). At the end of the calculations, we got the data of average response time and standard deviation (see Figure 4).

The average response time of 1.2 seconds indicates that NPC was generally quick to react to user inputs, providing a timely and interactive experience. This response time is acceptable for maintaining the flow of interactions in a virtual classroom. The standard deviation of 0.3 seconds shows that while most NPC responses were close to the average response time, there was some variability. This variability was mainly due to more complex interactions that required additional processing time. Three factors influence the response time value. The first is the complexity of interactions. Simple questions and commands resulted in faster response times, while complex or multi-step queries took longer for the NPC to process and respond. The second is AI processing capabilities. The underlying AI algorithms' efficiency and the VR platform's computational power significantly determined response times. Optimized algorithms and sufficient computational resources are crucial for minimizing delays. The third is network latency. In cases where interactions involved online components, network latency could affect response times. Ensuring a stable and fast network connection is essential for consistent performance.

The response time of AI-driven NPC in the educational simulation scenario was generally fast, with an average of 1.2 seconds and a standard deviation of 0.3 seconds. While most responses were timely, occasional delays in more complex interactions suggest areas for improvement in AI processing capabilities and computational resources. These findings provide valuable insights for optimizing AI performance in VR environments, ensuring smooth and responsive interactions that enhance the user experience.

```
using UnityEngine;
using System.Collections;
using System.IO;

public class ResponseTimeLogger : MonoBehaviour
{
    private string filePath = "response_times.csv";

    void Start()
    {
        if (!File.Exists(filePath))
        {
            File.WriteAllText(filePath, "UserInputTime,NPCResponseTime,ResponseTime\n")
        }
    }

    void LogResponseTime(float userInputTime, float npcResponseTime)
    {
        float responseTime = npcResponseTime - userInputTime;
        string logEntry = $"{userInputTime},{npcResponseTime},{responseTime}\n";
        File.AppendAllText(filePath, logEntry);
    }

    // Example function to be called on user input
    public void OnUserInput()
    {
        float userInputTime = Time.time;
        // Simulate NPC response time for demonstration purposes
        StartCoroutine(Respond(userInputTime));
    }

    IEnumerator Respond(float userInputTime)
    {
        // Simulate delay for NPC response
        yield return new WaitForSeconds(Random.Range(1.0f, 1.5f));
        float npcResponseTime = Time.time;
        LogResponseTime(userInputTime, npcResponseTime);
    }
}
```

Fig. 1. Logging script used to measure response time

```python
import pandas as pd

# Load the data
data = pd.read_csv('response_times.csv')

# Calculate average response time
average_response_time = data['ResponseTime'].mean()
standard_deviation = data['ResponseTime'].std()

print(f'Average Response Time: {average_response_time} seconds')
print(f'Standard Deviation: {standard_deviation} seconds')
```

Fig. 2. Script to calculate the response time in Python

```
UserInputTime,NPCResponseTime,ResponseTime
0.0,1.2,1.2
2.0,3.3,1.3
4.0,5.1,1.1
6.0,7.4,1.4
```

Fig. 3. Data and calculations in Python

```
Average Response Time: 1.25 seconds
Standard Deviation: 0.13 seconds
```

Fig. 4. The results of the average response time and standard deviation

The decision-making accuracy of AI-driven NPC in the educational simulation scenario was evaluated based on their ability to provide correct and contextually appropriate responses to user inputs. The accuracy rate was determined by comparing the NPC reactions to a set of predefined correct answers and assessing their relevance to the interaction context. The data is collected in three ways:

a) Participants interacted with NPC in a virtual classroom, where NPC acted as instructors and fellow students.
b) Two hundred interactions were recorded, where participants asked questions or made requests to NPCs, and the NPC responded.
c) Two independent evaluators reviewed each interaction and rated the NPC responses for accuracy and contextual relevance.

The decision-making accuracy of AI-driven NPC in the educational simulation scenario was evaluated using the following metrics. This overall metric is calculated by dividing the number of correct responses by the total number of interactions and multiplying by 100 to get a percentage.

a) Correct responses. It measures the number of responses provided by NPC that accurately address the questions or requests posed by users. Correct responses are factually accurate and contextually appropriate within the virtual classroom setting.

b) Incorrect responses. This metric counts the number of reactions that are factually incorrect or irrelevant to the user's question or request. Incorrect responses highlight areas where the AI's knowledge base or understanding needs improvement.
c) Ambiguous responses. This metric tracks the number of responses that are neither correct nor incorrect. Ambiguous responses lack sufficient clarity or context to be deemed accurate. These responses often result from the AI's difficulty processing complex or nuanced queries.
d) Accuracy rate. This overall metric is calculated by dividing the number of correct responses by the total number of interactions and multiplying by 100 to get a percentage.

Out of 200 interactions, 174 responses were rated as Correct, 18 as incorrect, and eight as ambiguous. The overall decision-making accuracy rate was calculated using the formula:

$$Accuracy\ rate = \frac{Number\ of\ correct\ responses}{Total\ number\ of\ interactions} \times 100\% \tag{1}$$

$$Accuracy\ rate = \frac{174}{200} \times 100\% = 87\%$$

From the data calculation, correct responses are 87%. Most NPC responses were accurate and contextually appropriate. These responses included providing correct answers to factual questions, giving relevant feedback on user actions, and appropriately engaging in discussions. For instance, when a user asked about the definition of photosynthesis, the NPC accurately explained the process, mentioning the role of sunlight, chlorophyll, and carbon dioxide in plant energy production. Incorrect responses are 9%. Some NPC responses were factually incorrect or did not address the user's question appropriately. For example, when asked about the capital of France, the NPC incorrectly responded with "Berlin." Such errors highlight areas where the AI's knowledge base or contextual understanding needs improvement. And ambiguous responses are 4%. A few responses were unclear or lacked sufficient context to be considered wholly accurate. These ambiguous responses often occurred in more complex interactions where the NPC failed to provide a definitive answer or the response was too vague. For example, when asked to explain a scientific concept, the NPC responded that it was too general and lacked specific details, making it difficult for the user to understand fully.

Three factors influence decision-making accuracy. The first is the complexity of queries. NPC performed better with straightforward questions compared to more complex or multi-part queries. It suggests that while the AI can handle fundamental interactions effectively, it struggles with more nuanced or detailed questions. The second is contextual understanding: The accuracy of NPC responses was higher in interactions where the context was clear and well-defined. Ambiguity in user inputs or complex contextual cues often led to less accurate responses, indicating a need for improved contextual processing in AI algorithms. The third is the knowledge Base. The scope and depth of the NPC knowledge base directly impacted their ability to provide accurate responses. Inaccurate or outdated information in the AI's knowledge base resulted in incorrect answers, highlighting the importance of maintaining a comprehensive and up-to-date knowledge repository.

Behavioral consistency in AI-driven NPC was assessed by evaluating how reliably these characters adhered to predefined rules and maintained consistent behavior across various interactions within the educational simulation scenario.

This evaluation is critical to ensure that NPC behaves predictably and realistically, enhancing the user's immersive experience in the virtual classroom. The flow to collect the data is divided into three ways:

a) One hundred fifty interactions between participants and NPCs were recorded in the virtual classroom.
b) The interactions were designed to cover a range of typical classroom activities, such as answering questions, providing feedback, and engaging in discussions.
c) Two independent evaluators reviewed each interaction to assess the consistency of NPC behavior.

Behavioral consistency was assessed using three metrics. These metrics collectively provide a comprehensive evaluation of NPC's ability to maintain consistent and realistic behavior, which is crucial for creating an immersive and believable virtual classroom environment:

a) Rule adherence. The degree to which NPC followed predefined rules governing their behavior in the virtual classroom. It evaluates how consistently NPC follow predefined regulations and guidelines governing their behavior. It ensures that NPCs act according to the established protocols for their roles in the virtual classroom.
b) Behavioral patterns. The consistency of NPC responses in similar situations across different interactions. It measures the consistency of NPC responses in similar situations across different interactions. It ensures that NPC exhibits predictable and reliable behavior when faced with identical user inputs or scenarios.
c) Interaction flow. The smoothness and coherence of NPC actions within ongoing interactions. It ensures NPC maintain a natural and continuous behavior flow throughout user interactions.

Non-player characters adhered to predefined behavioral rules in 138 out of 150 interactions, resulting in a rule adherence rate of 92%. This high rate indicates that NPC reliably followed the established guidelines for their roles in the virtual classroom. NPC generally followed predefined rules, such as providing specific feedback to students' questions and maintaining professional conduct. For example, when a student asked a question about a mathematical concept, the NPC consistently provided a structured explanation and relevant examples per the guidelines. Instances of rule deviation were minimal and often related to complex or unforeseen user inputs not covered by the rules.

NPCs exhibited consistent behavioral patterns in 135 out of 150 interactions, translating to a consistency rate of 90%. This measure reflects the NPC's ability to maintain predictable responses in similar contexts. NPC demonstrated consistent behavior patterns, such as using similar phrases and approaches when responding to similar questions. For instance, when students asked about the procedure for a scientific experiment, the NPC consistently outlined the steps clearly and systematically. Inconsistencies occurred occasionally when the NPC faced ambiguous or multi-part questions, leading to variations in their responses.

Evaluators rated 140 interactions as having smooth and coherent interaction flow, yielding a coherence rate of 93%. This metric captures the NPC's ability to engage in seamless and contextually appropriate actions throughout the interactions. Most interactions were rated as having a smooth and coherent flow. NPC managed to maintain the context and continuity of the conversation, even when switching between different topics or addressing multiple students. For example, NPC seamlessly transitioned between answering individual questions and facilitating group discussions during a class discussion. Minor disruptions in interaction flow were observed in scenarios involving complex multi-threaded conversations, where NPC sometimes struggled to manage concurrent interactions.

Three factors influencing behavioral consistency. First is the complexity of interactions. NPCs maintained higher consistency in straightforward interactions than more complex or multi-threaded conversations. It suggests that while AI algorithms effectively manage simple tasks, they may require further refinement to handle more intricate interactions seamlessly. The second is contextual variability. NPC performed consistently when the context of interactions was clear and well-defined. Variability in user inputs and the dynamic nature of the virtual classroom sometimes led to minor inconsistencies in NPC behavior, indicating a need for more robust contextual understanding capabilities. Last is algorithm robustness. The underlying AI algorithms played a significant role in ensuring behavioral consistency. Algorithms designed with robust rule-based frameworks and adaptive learning capabilities were more successful in maintaining consistent NPC behavior across different interactions.

The behavioral consistency of AI-driven NPCs in the educational simulation scenario was high, with rule adherence, behavioral patterns, and interaction flow rates of 92%, 90%, and 93%, respectively. These results indicate that NPCs generally maintained predictable and realistic behavior, contributing to a cohesive and immersive virtual classroom experience. However, areas for improvement were identified, particularly in handling complex interactions and enhancing contextual understanding. These findings provide valuable insights for developing more consistent and reliable AI-driven NPC for educational VR environments.

4.2 Factors affecting NPC performance

Several factors were identified in evaluating the performance of AI-driven NPC in the educational simulation scenario that significantly influence their behavior and interaction quality. Understanding these factors is crucial for improving AI algorithms and ensuring more realistic and practical NPC behavior in virtual classroom environments. Six factors that affect NPC performance include the complexity of user interaction, environmental context, AI algorithm limitations, computational resources, training data quality, and user adaptability.

The complexity of user interactions emerged as a significant factor affecting NPC performance. The NPC efficiently handled simple, straightforward questions or commands, resulting in quick and accurate responses. However, as the complexity of interactions increased—such as multi-part questions, ambiguous queries, or requests requiring detailed explanations—NPC performance showed noticeable variability.

a) Simple interactions. NPC responded effectively to direct questions or commands. These interactions had a high accuracy rate and consistent response times.
b) Complex interactions. In some cases, it involved more intricate questions. The decision-making accuracy for these complex interactions was lower, and response times were longer.

The virtual classroom's environmental context was crucial in influencing NPC performance. Factors such as the number of active participants, the type of classroom activities, and the overall classroom dynamics affected how NPC responded and interacted.

a) Number of participants. NPC maintained higher accuracy and consistency in sessions with fewer participants due to less cognitive load and fewer simultaneous interactions to manage. Conversely, in more crowded virtual classrooms, NPC sometimes exhibited delays and occasional inaccuracies as they processed multiple inputs.
b) Type of activities. Different classroom activities require varying levels of NPC engagement. For example, during lectures, NPC primarily needed to provide information and answer questions, which they managed effectively. However, during interactive group discussions or collaborative tasks, NPC faced more significant challenges in maintaining coherent and contextually appropriate behavior.

The limitations inherent in the AI algorithms used to control NPC behavior were a significant factor affecting performance. These limitations included NLP capabilities, contextual understanding, and decision-making algorithms.

a) The AI's ability to understand NLP inputs directly impacts NPC performance. While basic NLP tasks were handled well, more nuanced language processing—such as understanding idiomatic expressions, detecting sarcasm, or interpreting complex sentence structures—was less reliable. This limitation often led to incorrect or ambiguous responses.
b) Contextual understanding. Effective NPC performance requires a deep understanding of the context in which interactions occur. The AI algorithms sometimes struggle to maintain context across multi-turn conversations or understand the broader context of a user's question. This limitation resulted in less accurate and contextually relevant responses.
c) Decision-making algorithms. The algorithms governing NPC decision-making were generally practical for simple tasks but faced challenges with more complex decision processes. For instance, when multiple potential responses were equally valid, the AI occasionally selected suboptimal or less relevant options, affecting the perceived intelligence and reliability of the non-player character.

The availability and allocation of computational resources significantly impacted NPC performance. Adequate processing power, memory, and network capabilities ensured NPC could respond quickly and accurately to user inputs.

a) Processing power. NPC performance improved with greater processing power, allowing faster response times and more complex real-time decision-making. Limited processing resources led to slower responses and occasional performance bottlenecks.
b) Memory. Sufficient memory resources enabled NPC to maintain context over more extended interactions and handle more complex tasks without performance degradation. Memory constraints sometimes cause NPCs to lose context or simplify their responses.
c) Network latency. In scenarios involving online components or cloud-based AI processing, network latency was a critical factor. Low latency connections ensured smooth and timely interactions, while high latency could introduce delays and reduce the overall responsiveness of non-player character.

The quality and diversity of the training data used to develop AI algorithms were crucial for NPC performance. High-quality, comprehensive training data enabled the AI to learn various scenarios and respond appropriately.

a) Data diversity. Diverse training data covering various interaction types, user inputs, and contextual scenarios helped improve NPC performance. Insufficiently diverse data led to gaps in the AI's knowledge and reduced its ability to handle unfamiliar interactions.

b) Data accuracy. Accurate and well-annotated training data ensured NPC learned correct and contextually appropriate responses. Inaccurate or poorly annotated data could introduce biases or errors into the AI's decision-making processes, resulting in lower accuracy and consistency.

The ability of NPCs to adapt to individual user preferences and behaviors also affected their performance. Personalized interactions improved user satisfaction and engagement.

a) Personalisation. NPCs that adapted their responses based on user preferences, prior interactions, and learning styles were more effective in maintaining engagement and providing relevant information. Lack of personalization led to generic responses that might not fully address individual user needs.

b) User feedback integration. Incorporating real-time user feedback into NPC behavior adjustments helped improve interaction quality. NPC that could dynamically adjust based on immediate user reactions were perceived as more intelligent and responsive.

The performance of AI-driven NPC in the educational simulation scenario was influenced by multiple factors, including the complexity of user interactions, the environmental context, limitations of AI algorithms, computational resources, the quality of training data, and the ability to adapt to user preferences. Addressing these factors through improved AI algorithms, better resource allocation, diverse and accurate training data, and enhanced personalization techniques will be crucial for developing more effective and reliable NPC for virtual classroom environments.

4.3 User feedback

User feedback was collected through surveys and interviews, providing qualitative insights into the user experience with NPC in the educational simulation scenario (see Table 2). The overall satisfaction of participants interacting with AI-driven NPC in the educational simulation scenario was critical to this study. Participants were asked to complete surveys and participate in interviews after interacting with the NPC to gauge satisfaction. The feedback provided insights into various dimensions of user experience, including the realism, responsiveness, and engagement offered by the non-player character (see Table 3).

Table 2. Survey design and administration

No	Survey Design	Description
1	Survey Structure	The survey consisted of 5 Likert scale questions, ranging from 1 (strongly disagree) to 5 (strongly agree), covering different aspects of the user experience. The questions focused on NPC realism, responsiveness, engagement, usefulness, and overall satisfaction.
2	Participant Demographics	The study included 30 participants with varying levels of familiarity with VR technology, ensuring diverse perspectives.
3	Administration	Surveys were administered immediately after the interaction sessions to capture on-the-spot and accurate feedback. Follow-up interviews were conducted to delve deeper into specific aspects of user satisfaction.

Table 3. Results of the survey

No	Indicator	Score	Discussion
1	Realism	4.3	Participants rated the realism of NPC interactions highly, with an average score of 4.3 out of 5. Many participants noted that the NPC behavior and responses felt natural and lifelike, significantly enhancing the immersive experience.
2	Responsiveness	4.2	Users appreciated NPC's quick and timely responses, contributing to a seamless interaction flow. However, some participants mentioned occasional delays, particularly during complex queries.
3	Engagement	4.4	Participants felt that the NPC made the virtual classroom environment more interactive and engaging. The variety of interactions and the NPC's ability to maintain conversation context were highlighted as key strengths.
4	Usefulness	4.5	Participants found that NPC provided valuable information, guidance, and feedback, facilitating learning and understanding of the material.
5	Overall Satisfaction	4.2	Most participants expressed positive experiences, emphasising the potential of AI-driven NPC to enhance educational VR environments.

In addition to the quantitative survey data, qualitative insights from follow-up interviews provided a deeper understanding of user satisfaction (see Table 4).

Table 4. Interview insights

No	Item	Description
1	Positive Feedback	Participants frequently mentioned the NPCs' ability to create a realistic and engaging classroom atmosphere. The natural flow of interactions and the NPCs' ability to handle various questions and tasks were significant positives.
2	Areas for Improvement	Some participants pointed out areas where NPC performance could be enhanced. These included improving response times for complex queries, strengthening the depth of responses for more intricate topics, and reducing occasional inconsistencies in NPC behavior.
3	Impact on Learning	Many participants highlighted the positive impact of NPC on their learning experience. Interacting with lifelike virtual instructors and peers was a significant advantage, making learning more enjoyable and effective.

Participants' overall satisfaction with the AI-driven NPC in the educational simulation scenario was high, with an average satisfaction rating of 4.2 out of 5. The positive feedback centered on the realism, responsiveness, and engagement provided by the NPC, highlighting their potential to enhance the virtual classroom experience (see Table 5). However, the study also identified areas for improvement, particularly in handling complex queries and reducing response time variability. These insights provide valuable guidance for developing and optimizing AI-driven NPCs to meet user expectations and educational goals better [69].

Realism and engagement are crucial factors in assessing the effectiveness of AI-driven NPC in a virtual classroom setting. These factors determine how lifelike and immersive the interactions are, influencing the overall user experience and educational outcomes. The realism of NPC interactions was assessed based on user feedback and observational data. Participants were asked to rate the lifelikeness of NPC behavior, the naturalness of their responses, and the overall authenticity of their interactions.

Table 5. Realism survey results

No	Indicator	Score	Discussion
1	Natural Behavior	4.3	Users noted that NPC exhibited human-like behaviors, such as maintaining eye contact, using appropriate gestures, and modulating their tone of voice according to the interaction context. For instance, when explaining complex concepts, NPCs slowed their speech and used more detailed explanations, like how a human instructor would behave.
2	Response Authenticity	4.2	Participants felt that the NPC provided contextually appropriate and accurate responses, enhancing the interactions' realism. An example provided by a participant was an NPC explaining the process of photosynthesis in detail, including the role of sunlight, chlorophyll, and carbon dioxide, which matched their expectations of a knowledgeable instructor.
3	Behavioral Consistency	4.4	The consistency of NPC behavior was crucial in maintaining realism. Users observed that NPC consistently followed established interaction protocols.

Qualitative feedback has two critical points, including human-like interaction and seamless integration (see Table 6). Participants frequently highlighted the human-like quality of NPC interactions. Comments included praise for the NPC's ability to use facial expressions and body language to convey emotions and emphasis, making the interactions more genuine and engaging. Users appreciated the seamless integration of NPC into the virtual classroom environment. The NPC's ability to interact smoothly with the user and other NPC contributed to the overall sense of immersion. One participant mentioned that the NPC's ability to handle group discussions and manage classroom dynamics made the virtual environment like a real classroom.

Engagement measures how well the NPC captured and maintained user interest and participation throughout the interaction. High engagement levels indicate practical and interactive NPC design, essential for educational applications [70].

Table 6. Engagement survey results

No	Indicator	Score	Discussion
1	Interactive Experience	4.4	Participants found the virtual classroom more interactive than traditional learning methods due to NPC's dynamic and responsive nature. Users enjoyed the active participation and immediate feedback from NPC, which kept them engaged and motivated.
2	Attention Retention	4.5	The interactive elements, such as asking questions, providing prompts, and encouraging discussions, effectively kept users focused and involved. For example, NPC frequently checked in with students by asking follow-up questions or encouraging them to elaborate on their answers, which helped maintain engagement.
3	Motivation and Interest	4.3	Participants mentioned that the engaging nature of NPC interactions made them more interested in the subject matter and more willing to participate actively in virtual classroom activities.

Qualitative feedback has two crucial points, including interactive learning and engagement techniques. Participants praised the interactive learning experience facilitated by NPC. Many noted that the NPC's ability to simulate real-life classroom interactions, such as group discussions, Q&A sessions, and collaborative projects, made the learning process more dynamic and enjoyable. One user commented that they felt more involved and motivated to learn because the NPC made the lessons feel like a two-way conversation rather than a one-sided lecture. Users highlighted specific techniques NPC uses to enhance engagement, such as multimedia elements (e.g., virtual whiteboards, interactive diagrams) and gamification elements (e.g., quizzes, interactive challenges). These techniques were mentioned as effective ways to keep the sessions lively and engaging.

While realism and engagement were rated highly, some participants noted challenges in more complex interactions. NPC occasionally struggled with nuanced questions or multi-step problem-solving tasks, leading to less fluid interactions. Users suggested enhancing the AI's capability to handle complex queries more effectively to improve realism and engagement further. Participants expressed a desire for more personalized interactions. Although NPC performed well in general scenarios, adding more customized elements, such as remembering previous interactions and tailoring responses to individual learning styles, could enhance engagement [71]. One participant suggested implementing adaptive learning techniques to make NPC respond more specifically to their unique learning progress and preferences.

Artificial intelligence-driven NPC realism and engagement levels were high in the educational simulation scenario, contributing significantly to a positive user experience. Participants appreciated the NPC's lifelike behavior and authentic responses, which made the virtual classroom feel realistic and immersive. The high engagement scores reflected the NPC's ability to capture and maintain user interest through interactive and dynamic interactions. However, areas for improvement were identified, such as handling complex interactions and incorporating more personalized elements [72], [73], [74]. These insights provide valuable guidance for further development and optimiation of AI-driven NPC to enhance realism and engagement in virtual educational environments.

5 CONCLUSION

This study explored the accuracy and performance of AI-driven NPC within a VR classroom environment, focusing on an educational simulation scenario. The research aimed to provide comprehensive insights into the strengths and weaknesses of current AI implementations in VR settings by evaluating different AI techniques and their impact on NPC behavior.

The findings from the controlled experiments and user feedback revealed several key points. The NPC demonstrated a high decision-making accuracy rate of 87%, effectively handling user interactions and providing contextually appropriate responses, although there were occasional inaccuracies and ambiguities with complex queries. Behavioral consistency was strong, with NPC adhering to predefined rules and maintaining predictable behavior in 92% of interactions. However, the AI struggled with maintaining consistency in more dynamic and unpredictable scenarios. Response times were generally fast, averaging 1.2 seconds with a standard deviation of 0.3 seconds, though more complex interactions resulted in slightly slower responses.

Several factors influenced NPC performance. The complexity of user interactions significantly affected NPC performance, with more straightforward queries being handled more efficiently than complex or multi-step tasks. Environmental context, including the number of participants and type of activities, influenced NPC responsiveness and accuracy, with smaller, more focused sessions yielding better NPC performance. The inherent limitations of AI algorithms, such as NLP and contextual understanding, were identified as areas needing improvement to enhance NPC effectiveness.

Overall, user satisfaction with NPC interactions was high, with an average rating of 4.2 out of five. Participants appreciated NPC realism, responsiveness, and engagement, creating a more immersive and interactive learning experience. Realism and engagement were particularly praised, with NPCs rated highly for their lifelike behavior and ability to maintain user interest. However, participants also suggested improvements in handling complex interactions and personalizing responses.

Future development efforts should enhance AI capabilities to handle complex queries more effectively, particularly in NLP and contextual understanding. Optimizing computational resources for machine learning models can help balance performance and efficiency. Implementing adaptive learning algorithms that tailor NPC responses to individual user preferences and learning styles can further enhance engagement and educational outcomes. Adjustments can refine interactions and improve user satisfaction by incorporating real-time feedback into NPC behavior.

In conclusion, the study highlights the significant potential of AI-driven NPC to enhance the realism and interactivity of virtual educational environments. By addressing the identified areas for improvement, future developments can create more sophisticated and reliable NPCs, offering a more effective and engaging learning experience. The insights gained from this study provide a valuable foundation for advancing AI technologies in VR, contributing to the broader goal of creating immersive and impactful virtual experiences across various applications.

6 ACKNOWLEDGEMENTS

The author thanks the EQUITY (Enhancing Quality Education for International University Recognition) Funding Program for the International Research Collaboration 2024 from Universitas Negeri Yogyakarta under grant number 24.02.02/UN.34/SPK/PNLT-EQUITY/2024.

7 REFERENCES

[1] L. Zheng, T. Xie, and G. Liu, "Affordances of virtual reality for collaborative learning," in *Proc. – Int. Jt. Conf. Information, Media Eng. (ICIME),* 2018, pp. 6–10. https://doi.org/10.1109/ICIME.2018.00011

[2] S. M. Reeves and K. J. Crippen, "Virtual laboratories in undergraduate science and engineering courses: A systematic review, 2009–2019," *J. Sci. Educ. Technol.*, vol. 30, pp. 16–30, 2021. https://doi.org/10.1007/s10956-020-09866-0

[3] A. Scavarelli, A. Arya, and R. J. Teather, "Virtual reality and augmented reality in social learning spaces: A literature review," *Virtual Real.*, vol. 25, pp. 257–277, 2021. https://doi.org/10.1007/s10055-020-00444-8

[4] J. M. Loomis, J. J. Blascovich, and A. C. Beall, "Immersive virtual environment technology as a basic research tool in psychology," *Behav. Res. Methods, Instruments, Comput.*, vol. 31, pp. 557–564, 1999. https://doi.org/10.3758/BF03200735

[5] S. Feng and A. H. Tan, "Towards autonomous behavior learning of non-player characters in games," *Expert Syst. Appl.*, vol. 56, pp. 89–99, 2016. https://doi.org/10.1016/j.eswa.2016.02.043

[6] H. J. Yoon, J. Kim, S. W. Park, and H. Heo, "Influence of virtual reality on visual parameters: Immersive versus non-immersive mode," *BMC Ophthalmol.*, vol. 20, p. 200, 2020. https://doi.org/10.1186/s12886-020-01471-4

[7] W. Xu, K. Yu, X. Meng, D. Monteiro, D. Kao, and H. N. Liang, "Exploring the effect of the Group Size and Feedback of non-player character spectators in virtual reality exergames," *Front. Psychol.*, vol. 14, p. 1079132, 2023. https://doi.org/10.3389/fpsyg.2023.1079132

[8] P. Athanasiou, E. Voyiatzaki, and I. Hatzilygeroudis, "Evolving non-player characters in educational games in virtual worlds," *EDULEARN23 Proc.*, 2023, pp. 7053–7059. https://doi.org/10.21125/edulearn.2023.1851

[9] J. Burgess and C. Jones, "'I harbour strong feelings for Tali despite her being a fictional character': Investigating videogame players' emotional attachments to non-player characters," *Game Stud.*, vol. 20, no. 1, 2020. https://gamestudies.org/2001/articles/burgessjones

[10] S. O'Connor, N. Gasteiger, E. Stanmore, D. C. Wong, and J. J. Lee, "Artificial intelligence for falls management in older adult care: A scoping review of nurses' role," *J. Nurs. Manag.*, vol. 30, no. 8, pp. 3787–3801, 2022. https://doi.org/10.1111/jonm.13853

[11] L. Jensen and F. Konradsen, "A review of the use of virtual reality head-mounted displays in education and training," *Educ. Inf. Technol.*, vol. 23, pp. 1515–1529, 2018. https://doi.org/10.1007/s10639-017-9676-0

[12] L. Maffei, M. Masullo, A. Pascale, G. Ruggiero, and V. Puyana Romero, "On the validity of immersive virtual reality as tool for multisensory evaluation of urban spaces," *Energy Procedia*, vol. 78, pp. 471–476, 2015. https://doi.org/10.1016/j.egypro.2015.11.703

[13] R. P. McMahan, D. A. Bowman, D. J. Zielinski, and R. B. Brady, "Evaluating display fidelity and interaction fidelity in a virtual reality game," *IEEE Trans. Vis. Comput. Graph.*, vol. 18, no. 4, pp. 626–633, 2012. https://doi.org/10.1109/TVCG.2012.43

[14] C. Figueiredo, F. Rebelo, P. Noriega, and E. Vilar, "Storyboards in VR narratives planning: How to create and evaluate them," in *HCI International 2022 – Late Breaking Papers: Interacting with eXtended Reality and Artificial Intelligence, HCII 2022,* in Lecture Notes in Computer Science, J. Y. C. Chen, G. Fragomeni, H. Degen, and S. Ntoa, Eds., Springer, Cham, vol. 13518, 2022, pp. 13–32. https://doi.org/10.1007/978-3-031-21707-4_2

[15] J. W. V. De Faria, M. J. Teixeira, L. De Moura Sousa Júnior, J. P. Otoch, and E. G. Figueiredo, "Virtual and stereoscopic anatomy: When virtual reality meets medical education," *J. Neurosurg.*, vol. 125, no. 5, pp. 1105–1111, 2016. https://doi.org/10.3171/2015.8.JNS141563

[16] J. Stenros, "Between game facilitation and performance: Interactive actors and non-player characters in Larps," *Int. J. Role-Playing*, no. 4, pp. 76–93, 2023. https://doi.org/10.33063/ijrp.vi4.230

[17] J. Moon, "Reviews of social embodiment for design of non-player characters in virtual reality-based social skill training for autistic children," *Multimodal Technol. Interact.*, vol. 2, no. 3, p. 53, 2018. https://doi.org/10.3390/mti2030053

[18] H. Warpefelt and H. Verhagen, "A model of non-player character believability," *J. Gaming Virtual Worlds*, vol. 9, no. 1, pp. 39–53, 2017. https://doi.org/10.1386/jgvw.9.1.39_1

[19] S. Lapeyrade and C. Rey, "Non-player character decision-making with prolog and ontologies," in *IEEE Conf. Comput. Intell. Games (CoG)*, 2023, pp. 1–2. https://doi.org/10.1109/CoG57401.2023.10333221

[20] Y. H. Lai, S. Y. Chen, C. F. Lai, Y. C. Chang, and Y. S. Su, "Study on enhancing AIoT computational thinking skills by plot image-based VR," *Interact. Learn. Environ.*, vol. 29, no. 3, pp. 482–495, 2021. https://doi.org/10.1080/10494820.2019.1580750

[21] E. A. L. Lee and K. W. Wong, "Learning with desktop virtual reality: Low spatial ability learners are more positively affected," *Comput. Educ.*, vol. 79, pp. 49–58, 2014. https://doi.org/10.1016/j.compedu.2014.07.010

[22] F. Ke, M. Pachman, and Z. Dai, "Investigating educational affordances of virtual reality for simulation-based teaching training with graduate teaching assistants," *J. Comput. High. Educ.*, vol. 32, pp. 607–627, 2020. https://doi.org/10.1007/s12528-020-09249-9

[23] D. Hamilton, J. McKechnie, E. Edgerton, and C. Wilson, "Immersive virtual reality as a pedagogical tool in education: A systematic literature review of quantitative learning outcomes and experimental design," *J. Comput. Educ.,* vol. 8, pp. 1–32, 2021. https://doi.org/10.1007/s40692-020-00169-2

[24] M. Lorenz *et al.*, "Presence and user experience in a virtual environment under the influence of ethanol: An explorative study," *Sci. Rep.*, vol. 8, p. 6407, 2018. https://doi.org/10.1038/s41598-018-24453-5

[25] X. Pan and A. F. d. C. Hamilton, "Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape," *Br. J. Psychol.*, vol. 109, no. 3, pp. 395–417, 2018. https://doi.org/10.1111/bjop.12290

[26] M. Tomasz and G. Michael, "Virtual reality history, applications, technology and future," Technical Report, TR-186-2-96-06, 1996. [Online]. Available: https://www.cg.tuwien.ac.at/research/publications/1996/mazuryk-1996-VRH/mazuryk-1996-VRH-paper.pdf

[27] Y. Shu and T. C. Huang, "Identifying the potential roles of virtual reality and STEM in Maker education," *J. Educ. Res.*, vol. 114, no. 2, pp. 108–118, 2021. https://doi.org/10.1080/00220671.2021.1887067

[28] D. Allcoat and A. von Mühlenen, "Learning in virtual reality: Effects on performance, emotion and engagement," *Res. Learn. Technol.*, vol. 26, 2018. https://doi.org/10.25304/rlt.v26.2140

[29] J. Moloney, B. Spehar, A. Globa, and R. Wang, "The affordance of virtual reality to enable the sensory representation of multi-dimensional data for immersive analytics: From experience to insight," *J. Big Data*, vol. 5, 2018. https://doi.org/10.1186/s40537-018-0158-z

[30] C. Wienrich and M. E. Latoschik, "eXtended artificial intelligence: New prospects of human-AI interaction research," *Front. Virtual Real.*, vol. 2, p. 686783, 2021. https://doi.org/10.3389/frvir.2021.686783

[31] O. Ali, P. A. Murray, M. Momin, Y. K. Dwivedi, and T. Malik, "The effects of artificial intelligence applications in educational settings: Challenges and strategies," *Technol. Forecast. Soc. Change*, vol. 199, p. 123076, 2024. https://doi.org/10.1016/j.techfore.2023.123076

[32] A. Palmquist, I. Jedel, C. Hart, and A. Soellaart, "'Not in Kansas anymore' Exploring Avatar-Player dynamics through a Wizard of Oz approach in virtual reality," in *Virtual, Augmented and Mixed Reality: 16th International Conference, VAMR 2024*, 2024, pp. 259–276. https://doi.org/10.1007/978-3-031-61041-7_17

[33] P. García-Sánchez, "Georgios N. Yannakakis and Julian Togelius: Artificial intelligence and games," *Genet. Program. Evolvable Mach.*, vol. 20, pp. 143–145, 2019. https://doi.org/10.1007/s10710-018-9337-0

[34] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015. https://doi.org/10.1038/nature14539

[35] R. K. Mishra, G. Y. S. Reddy, and H. Pathak, "The understanding of deep learning: A comprehensive review," *Math. Probl. Eng.*, vol. 2021, no. 1, pp. 1–15, 2021. https://doi.org/10.1155/2021/5548884

[36] N. Silaparasetty, "Machine learning vs. Deep learning," in *Mach. Learn. Concepts with Python Jupyter Noteb. Environ.*, Berkeley, CA: Apress, 2020, pp. 57–65. https://doi.org/10.1007/978-1-4842-5967-2_4

[37] W. Pedrycz and S. M. Chen, *Deep Learning: Algorithms and Applications Deep Learning*. Cham: Springer, 2020. https://doi.org/10.1007/978-3-030-31760-7

[38] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, pp. 529–533, 2015. https://doi.org/10.1038/nature14236

[39] Y. Kang and A. H. Tan, “Self-organising agents for reinforcement learning in virtual worlds,” in *Proc. Int. Jt. Conf. Neural Networks (IJCNN)*, 2010, pp. 1–8. https://doi.org/10.1109/IJCNN.2010.5596363

[40] Y. Ye, X. Zhang, and J. Sun, “Automated vehicle’s behavior decision making using deep reinforcement learning and high-fidelity simulation environment,” *Transp. Res. Part C Emerg. Technol.*, vol. 107, pp. 155–170, 2019. https://doi.org/10.1016/j.trc.2019.08.011

[41] D. Jurafsky and J. H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, Second Ed., Upper Saddle River, NJ: Pearson, 2009. [Online]. Available: http://www.cs.colorado.edu/~martin/slp.html

[42] A. Vaswani *et al.*, “Retention is all you need,” in *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM ’23)*, 2023, pp. 4752–4758. https://doi.org/10.1145/3583780.3615497

[43] J. Togelius, M. Preuss, and G. N. Yannakakis, “Towards multiobjective procedural map generation,” in *Proceedings of the 2010 Workshop on Procedural Content Generation in Games (PCGames ’10)*, 2010, pp. 1–8. https://doi.org/10.1145/1814256.1814259

[44] J. Togelius *et al.*, “Controllable procedural map generation via multiobjective evolution,” *Genet. Program. Evolvable Mach.*, vol. 14, pp. 245–277, 2013. https://doi.org/10.1007/s10710-012-9174-5

[45] J. Togelius, G. N. Yannakakis, K. O. Stanley, and C. Browne, “Search-based procedural content generation: A taxonomy and survey,” *IEEE Trans. Comput. Intell. AI in Games*, vol. 3, no. 3, pp. 172–186, 2011. https://doi.org/10.1109/TCIAIG.2011.2148116

[46] J. Blascovich, J. Loomis, A. C. Beall, K. R. Swinth, C. L. Hoyt, and J. N. Bailenson, “Immersive virtual environment technology: Just another methodological tool for social psychology?” *Psychol. Inq.*, vol. 13, no. 2, pp. 146–149, 2002. https://doi.org/10.1207/S15327965PLI1302_03

[47] R. Calo and K. Crawford, “There is a blind spot in AI research,” *Nature*, vol. 538, pp. 311–313, 2016. https://doi.org/10.1038/538311a

[48] M. Slater *et al.*, “The ethics of realism in virtual and augmented reality,” *Front. Virtual Real.*, vol. 1, 2020. https://doi.org/10.3389/frvir.2020.00001

[49] M. Hofer, T. Hartmann, A. Eden, R. Ratan, and L. Hahn, “The role of plausibility in the experience of spatial presence in virtual environments,” *Front. Virtual Real.*, vol. 1, 2020. https://doi.org/10.3389/frvir.2020.00002

[50] M. E. Latoschik and C. Wienrich, “Congruence and plausibility, not presence: Pivotal conditions for XR experiences and effects, a novel approach,” *Front. Virtual Real.*, vol. 3, 2022. https://doi.org/10.3389/frvir.2022.694433

[51] N. Miketić, I. Pinćjer, and A. Lilić, “Integration of the visual elements of art and personality factors in process of character design,” University of Novi Sad, Faculty of Technical Sciences, Department of Graphic Engineering and Design, pp. 533–539, 2018. https://doi.org/10.24867/GRID-2018-p64

[52] M. Mateas and A. Stern, “Structuring content in the façade interactive drama architecture,” *Proceedings of the First Artificial Intelligence and Interactive Digital Entertainment Conference*, vol. 1, no. 1, pp. 93–98, 2021. https://doi.org/10.1609/aiide.v1i1.18722

[53] M. Mateas and A. Stern, “Interaction and Narration,” in *The Antipodes of the Mind: Charting the Phenomenology of the Ayahuasca Experience.* Oxford: Oxford Academic, 2023, pp. 99–112. https://doi.org/10.1093/oso/9780199252923.003.0008

[54] A. Christopoulos, M. Conrad, and M. Shukla, “Increasing student engagement through virtual interactions: How?” *Virtual Real.*, vol. 22, pp. 353–369, 2018. https://doi.org/10.1007/s10055-017-0330-3

[55] P. Thagard, J. Gratch, and S. Marsella, “A domain-independent framework for modeling emotion,” *Cogn. Syst. Res.*, vol. 5, no. 4, pp. 269–306, 2004. https://doi.org/10.1016/j.cogsys.2004.02.002

[56] J. E. Laird and M. Van Lent, "Human-level AI's killer application interactive computer games," *AI Mag.*, vol. 22, no. 2, pp. 15–25, 2001.

[57] M. O. Riedl and A. Stern, "Believable agents and intelligent story adaptation for interactive storytelling," in *Technologies for Interactive Digital Storytelling and Entertainment, TIDSE 2006,* in Lecture Notes in Computer Science, S. Göbel, R. Malkewitz, and I. Iurgel, Eds., vol. 4326, Springer, Berlin, Heidelberg, 2006, pp. 1–12. https://doi.org/10.1007/11944577_1

[58] S. Ventura, E. Brivio, G. Riva, and R. M. Baños, "Immersive versus non-immersive experience: Exploring the feasibility of memory assessment through 360° technology," *Front. Psychol.*, vol. 10, 2019. https://doi.org/10.3389/fpsyg.2019.02509

[59] B. Liu and W. Zhou, "Research on China's AR/VR industry transformation strategies in the context of domestic and foreign market changes," *E3S Web of Conf.*, vol. 235, 2021, pp. 1–4. https://doi.org/10.1051/e3sconf/202123503037

[60] T. Inamura and Y. Mizuchi, "SIGVerse: A cloud-based VR platform for research on multimodal human-robot interaction," *Front. Robot. AI*, vol. 8, 2021. https://doi.org/10.3389/frobt.2021.549360

[61] L. Floridi *et al.*, "AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations," *Minds & Mach.*, vol. 28, pp. 689–707, 2018. https://doi.org/10.1007/s11023-018-9482-5

[62] A. Nguyen, H. N. Ngo, Y. Hong, B. Dang, and B. P. T. Nguyen, "Ethical principles for artificial intelligence in education," *Educ. Inf. Technol.*, vol. 28, pp. 4221–4241, 2023. https://doi.org/10.1007/s10639-022-11316-w

[63] G. Davidson, G. Todd, J. Togelius, T. M. Gureckis, and B. M. Lake, "Goals as reward-producing programs," *arXiv preprint arXiv:2405.13242*, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2405.13242

[64] L. Daniela and M. D. Lytras, "Editorial: Themed issue on enhanced educational experience in virtual and augmented reality," *Virtual Real.*, vol. 23, pp. 325–327, 2019. https://doi.org/10.1007/s10055-019-00383-z

[65] W. S. Alhalabi, "Virtual reality systems enhance students' achievements in engineering education," *Behav. Inf. Technol.*, vol. 35, no. 11, pp. 919–925, 2016. https://doi.org/10.1080/0144929X.2016.1212931

[66] J. Loh, "Inquiry into issues of trustworthiness and quality in narrative studies: A perspective," *The Qual. Rep.*, vol. 18, no. 33, pp. 1–5, 2013. https://doi.org/10.46743/2160-3715/2013.1477

[67] A. A. Rafiq *et al.*, "Research Instrument – In-World NPCs: Analysing artificial intelligence precision in virtual reality settings," Zenodo, 2024. https://doi.org/10.5281/zenodo.13118065

[68] M. K. W. Stranges, S. Ul Haq, and D. G. Dunn, "Black-out test versus UV camera for corona inspection of HV motor stator endwindings," *IEEE Trans. Ind. Appl.*, vol. 50, no. 5, pp. 3135–3140, 2014. https://doi.org/10.1109/TIA.2014.2306979

[69] A. Das, S. Malaviya, and M. Singh, "The impact of AI-driven personalization on learners' performance," *Artic. Int. J. Comput. Sci. Eng.*, vol. 11, no. 8, pp. 15–22, 2023. https://www.ijcseonline.org/pub_paper/3-IJCSE-09211.pdf

[70] H. A. El-Sabagh, "Adaptive e-learning environment based on learning styles and its impact on development students' engagement," *Int. J. Educ. Technol. High. Educ.*, vol. 18, 2021. https://doi.org/10.1186/s41239-021-00289-4

[71] A. Bhagat, R. Vyas, and T. Singh, "Students awareness of learning styles and their perceptions to a mixed method approach for learning," *Int. J. Appl. Basic Med. Res.*, vol. 5, no. Suppl. 1, pp. S58–S65, 2015. https://doi.org/10.4103/2229-516X.162281

[72] Y. Xu *et al.*, "Artificial intelligence: A powerful paradigm for scientific research," *The Innovation*, vol. 2, no. 4, 2021. https://doi.org/10.1016/j.xinn.2021.100179

[73] M. Bilal, Y. Zhang, S. Cai, U. Akram, and A. Halibas, “Artificial intelligence is the magic wand making customer-centric a reality! An investigation into the relationship between consumer purchase intention and consumer engagement through affective attachment,” *J. Retail. Consum. Serv.*, vol. 77, p. 103674, 2024. https://doi.org/10.1016/j.jretconser.2023.103674

[74] W. Xu and F. Ouyang, “The application of AI technologies in STEM education: A systematic review from 2011 to 2021,” *Int. J. STEM Educ.*, vol. 9, 2022. https://doi.org/10.1186/s40594-022-00377-5

8 AUTHORS

Mochamad Bruri Triyono is a Professor of the TVET (Technical and Vocational Education and Training) Department at the Graduate School of the Universitas Negeri Yogyakarta, Indonesia. He is also a Lecture in the department of Mechanical Engineering Education at the Universitas Negeri Yogyakarta, Indonesia. His main research interests include TVET and instructional media (E-mail: bruritriyono@uny.ac.id).

Arif Ainur Rafiq is a Lecture in the Department of Mechatronics Engineering at the Politeknik Negeri Cilacap, Indonesia. His main research interests include sensors and transducers, robotics, embedded systems, TVET, instructional media, augmented and virtual reality for education, and artificial intelligence (E-mail: aar@pnc.ac.id).

Didik Hariyanto is an Associate Professor and researcher at the Department of Electrical Engineering Education, Faculty of Engineering, Universitas Negeri Yogyakarta, Indonesia. He earned his Doctoral degree from the Institute of Vocational Education and Vocational Didactics, TU Dresden, Germany. Presently, he serves as the Head of the Center of STEAM and Advanced Learning Technology at Universitas Negeri Yogyakarta, Indonesia. His research focuses on adaptive e-learning, technology-enhanced learning, and education within the realm of electrical engineering (E-mail: didik_hr@staff.uny.ac.id).

Dina Adinda Dina Adinda is an Associate Professor and researcher in education and training sciences at Université Paris Nanterre, France. She holds a Ph.D. in Education and Training and a master’s degree in Instructional Design and Technology both from University of Strasbourg, France. Her research interests include higher education, adults learning, students’ autonomy, self-direction, and technology for teaching and learning (E-mail: dadinda@parisnanterre.fr).

Maria Denami Maria Denami is an Associate Professor of Education Sciences. Interested in Higher Education Pedagogy, her main research topic is the process of professionalization of students and their support to the development of professional skills. In this sense, her research objective is to understand how experiential learning combined with reflexivity support leads to the awareness/development of students’ professional competences. She is a member of LISEC – UR 2310. She has worked for three years in the Student Success Unit and has designed and managed various projects to support integration and students’ success (E-mail: maria.denami@uha.fr).