

System Development and Evaluation for Mass Casualty Incidents Triage with Virtual Reality and Artificial Intelligence

Peng Xia

Auckland University of Technology
peng.xia@aut.ac.nz

Ji Ruan *

Auckland University of Technology
ji.ruan@aut.ac.nz

David Parry

Murdoch University
david.parry@murdoch.edu.au

Stephen Aiello

Auckland University of Technology
stephen.aiello@aut.ac.nz

Jian Yu

Auckland University of Technology
jian.yu@aut.ac.nz

Sally Britnell

Auckland University of Technology
sally.britnell@aut.ac.nz

ABSTRACT

This study investigates the integration of Virtual Reality (VR) and Artificial Intelligence (AI) to enhance pre-hospital triage training for Mass Casualty Incidents (MCIs). Traditional training methods, such as field drills and full-scale simulations, are often costly and logistically challenging, while simpler methods like tabletop exercises remain limited in realism. To address these limitations, a VR learning tool was developed to simulate realistic emergency scenarios, providing emergency healthcare professionals with an immersive and cost-effective training environment to refine triage skills. The VR learning tool records both VR sensor data and speech data, and then utilizes statistical and AI methods (such as automatic speech recognition, and natural language processing) to process these data for evaluation. The survey results showed that participants with varying levels of experience found the VR training highly immersive and engaging. Additionally, AI-driven analysis of speech data from the training demonstrated improved consistency and correctness in participants' communication over time. This research demonstrates VR's potential as a valuable supplement to traditional training, identifying key areas for future development.

Keywords

Mass Casualty Incidents, Virtual Reality, Artificial Intelligence, Education, Training

INTRODUCTION

Mass Casualty Incidents (MCIs), as a result of natural or human-made disasters, require emergency personnel to make critical decisions quickly under extreme stress (Marx et al., 2013). Traditional training methods, such as field drills and full-scale simulations, are often costly and logistically challenging, while simpler methods like tabletop exercises remain limited in realism. Virtual Reality (VR) provides an effective alternative, allowing realistic and immersive training scenarios at a reduced cost compared to live simulations (Mills et al., 2020). Current research has explored the role of VR in emergency scenarios such as car accidents (Berndt, Wessel, Mentler, and Herczeg, 2018; Koutitas et al., 2019) and natural disasters (Foronda et al., 2016; Ingrassia et al., 2015). Meanwhile, Artificial Intelligence (AI) research has advanced resource optimization and assessment in emergency departments (ED),

*corresponding author

focusing on tasks such as predicting casualty and reducing waiting time (Billis et al., 2019; Cabrera et al., 2012; Caicedo-Torres et al., 2016).

Despite these advancements, notable gaps persist in the integration of VR and AI for MCI training. Previous VR-based studies primarily rely on the evaluation methods such as questionnaires and interviews (Lineberry et al., 2013; Maclure, 2021), which can be subjective and fail to utilize objective data from advanced sensors. Additionally, while AI applications in emergency healthcare predominantly focus on in-hospital scenarios, there is limited research addressing pre-hospital triage for MCIs (Fan, 2012; Tahayori et al., 2021; Yan et al., 2019). Furthermore, advanced AI techniques like speech and image recognition are underutilized in this domain, leaving opportunities for more sophisticated applications.

This project introduces a VR learning tool that integrates AI, tailored specifically for pre-hospital triage training for paramedic students responding to MCIs. While MCIs may involve multiple emergency services such as police or fire personnel, the scope of this study is focused exclusively on the training needs of paramedics, who will play a critical frontline role in casualty assessment and communication during such events. By creating an immersive and realistic training environment, the VR learning tool aims to overcome the financial and logistical barriers of traditional methods. It facilitates paramedic students, especially in the pre-hospital stage, in practising casualty assessment, triage, and critical decision-making in a controlled, risk-free setting, while also enabling participants to practice verbal communication such as the METHANE (Major incident, Exact location, Type of incident, Hazards, Access and egress, Number of casualties and Emergency services) report. The METHANE report is a communication protocol which is used to provide critical information during major incidents, including details on the situation, hazards, casualties, and required resources. The tool leverages AI and Machine Learning (ML) to analyze diverse data types such as time span, audio speech and decision-making processes, providing detailed insights into training outcomes and areas for improvement.

While a limited number of studies have explored VR applications for pre-hospital triage, these often focus on observational training, lack interactive engagement with virtual objects, or rely primarily on subjective evaluations. In contrast, this study introduces a more comprehensive triage training and evaluation system that integrates sensor and speech data capture with AI-driven performance assessment, in addition to the more traditional questionnaire-based evaluations. The system captures real-time physical interactions with virtual casualties and evaluates verbal communication skills using AI models, providing a multi-layered, data-rich understanding of participant competence. This approach marks a methodological advancement by introducing objective and scalable assessment mechanisms into pre-hospital triage training.

This research is guided by the following key questions:

1. What are the essential components required to build an effective VR learning tool?
2. How can AI be integrated effectively into a VR learning tool to enhance data analysis?
3. How can the training effectiveness of the VR learning tool be evaluated and improved?

The first two questions focus on the foundational aspects of developing and enhancing the VR learning tool. Specifically, they address the identification of essential components—both hardware and software—that contribute to effective training, ensuring a robust and functional VR training environment. Additionally, these questions examine the integration of AI into the system, emphasizing how advanced techniques like visual and voice recognition can enrich data analysis, providing deeper insights into participants' performance and outcomes.

The third question focuses on evaluating and improving the training effectiveness of the VR learning tool. This includes exploring factors such as correctness and training duration to refine the system for more efficient and impactful learning experiences. The proposed sub-questions include: (1) *How can data from the VR learning tool be used to assess the competence of participants in MCI scenarios?* (2) *Can a VR environment convince participants to behave in a way that will allow their performance to be assessed?* and (3) *How well do participants accept VR technology for triage training in MCIs?*

To address these questions, the research is structured into several phases: defining relevant training tasks, constructing a data collection platform with appropriate hardware (e.g., VR headsets, controllers, and sensors) and software for simulated scenarios, conducting experiments to capture diverse data types such as task completion times, speech, and interactions and analyzing the collected data using ML and traditional statistical methods. This structured approach aims to ensure the VR learning tool's effectiveness in improving training for MCIs.

LITERATURE REVIEW

The literature review of this study provides a comprehensive synthesis of key advancements in VR and AI applications in healthcare, with a specific focus on emergency training, triage, and decision-making processes. These insights highlight the transformative potential of VR and AI technologies while identifying critical gaps that warrant future exploration.

Applications of VR in Healthcare Training and Education

VR has established itself as a versatile tool for healthcare training, offering risk-free environments for skill development. Early research emphasized its potential to simulate complex scenarios for enhanced engagement and learning outcomes. Mantovani et al. (2003) and Gaba (2004) demonstrated VR's utility in critical care training, highlighting its ability to provide realistic simulations without endangering patients. Aggarwal et al. (2006) showed VR's efficacy in laparoscopic skill acquisition, which was validated by Larsen et al. (2009), who confirmed effective skill transfer from VR to clinical settings.

In nursing and anatomy education, Kilmon et al. (2010) and Izard and Méndez (2016) explored VR's potential to improve learning precision and interactive engagement. These applications underline VR's role in both technical skill development and the enhancement of soft skills such as decision-making and patient communication. More recent surveys and reviews by Mistry et al. (2023), Guo et al. (2024), and Sung et al. (2024) show the growing popularity and importance of VR in healthcare training and education.

VR and AI for Emergency Healthcare and Triage

Andreatta et al. (2010) and Cone et al. (2011) showed that VR enhances situational awareness and decision-making in emergency scenarios. Ingrassia et al. (2015) validated VR-based triage training, while Berndt, Wessel, Mentler, and Herczeg (2018) and Koutitas et al. (2019) highlighted its adaptability for diverse applications. Additionally, Berndt, Wessel, Willer, et al. (2018) further confirmed VR's immersion and presence in MCI training.

AI has complemented these efforts by optimizing decision-making in triage systems. Yan et al. (2019) and Kang et al. (2020) leveraged ML algorithms to enhance triage efficiency, while Kerr et al. (2021) utilized fuzzy logic in robotic systems to manage ambiguous patient data. Billis et al. (2019) and Cabrera et al. (2012) focused on optimizing emergency department (ED) workflows through resource allocation models. These studies collectively highlight AI's capability to streamline emergency care operations and improve patient outcomes.

More recent studies reaffirm VR and AI's potential while identifying limitations. Harada et al. (2024) and Behmadi et al. (2022) showed VR's impact on The Simple Triage And Rapid Treatment (START) triage training, emphasizing skill retention and user engagement. Paslı et al. (2024) evaluated GPT-4's triage accuracy, demonstrating alignment with in-hospital physician decisions but highlighting interpretability challenges. While Paslı et al. (2024) demonstrated GPT-4's potential for predicting triage outcomes in emergency departments, the present study applies GPT-4 in a pre-hospital environment to evaluate structured verbal communication using the METHANE communication protocol. Rather than scoring diagnostic accuracy, our implementation focuses on semantic completeness, offering a different application of GPT-4 in evaluating speech data from the training sessions.

METHODOLOGY

This study used design science methodology and employed a comprehensive research framework to develop and evaluate a VR learning tool for MCIs training. The framework consisted of three major components (see Figure 1): development tools, the VR Learning Tool, and the VR training experiment.

Component 1: Development Tools

To develop an effective and immersive VR learning tool, the project employed two categories of development tools: VR development tools and AI development tools. Each played a vital role in creating a realistic VR environment and integrating advanced AI capabilities for data analysis and interactivity. The VR development tools included modelling software and the Unity Engine, which facilitated the creation and assembly of 3D objects and virtual environments. As shown in Figure 2, the modelling process followed a structured workflow, starting with the identification of necessary objects such as vehicles, buildings, and humanoid characters. These objects were categorized as static or dynamic based on their role in the simulation. Static objects, like cars and streetlights, were created using Blender and Google SketchUp, while dynamic objects, including humanoid casualties, underwent a multi-step polygon modelling process. For simple characters, low polygon modelling was used with Adobe Fusion

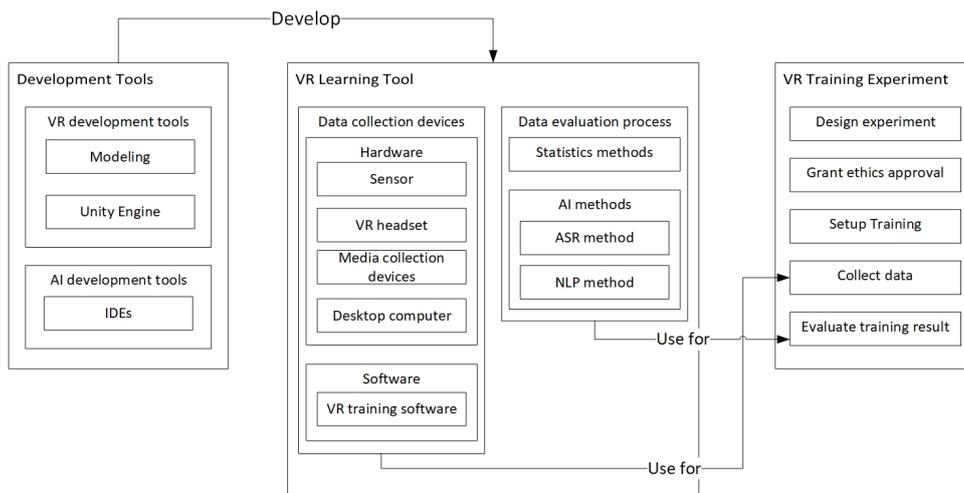


Figure 1. Research Framework

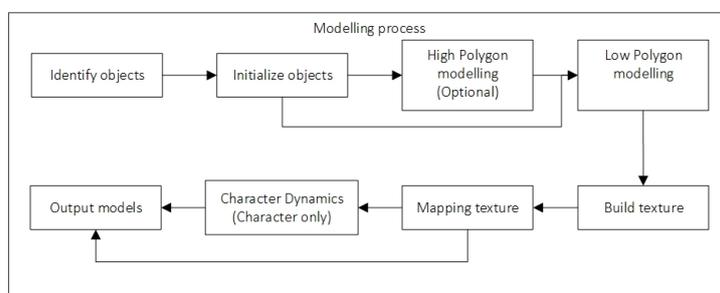


Figure 2. Virtual Character Modeling Workflow

whereas more complex characters requiring detailed features, such as trauma injuries, were first crafted in high polygon format using ZBrush or Maya and then optimized to lower polygons for performance efficiency. Textures, including those representing medical conditions, were designed in Adobe Photoshop and integrated into the models using Maya or Unity, adding realism to the virtual casualties. Animated humanoid characters were equipped with dynamic joint structures, with approximately 20 joint groups per character, enabling natural movement and interaction. The final models were exported as *.FBX files and imported into the Unity Engine for integration into the VR scenarios.

The Unity Engine served as the central platform for assembling and scripting the VR environment. Figure 3 shows the entire workflow in the Unity Engine. The development process began with constructing a detailed urban landscape, combining elements from the Unity asset marketplace with custom designs. Static objects, such as vehicles, rocks, and roads, were strategically placed to provide context and visual realism to the scenarios. Dynamic objects, such as virtual casualties, were programmed with detailed interaction behaviours. These interactions were defined through scripting in Microsoft Visual Studio, enabling participants to engage with the environment realistically. For example, actions such as managing a casualty’s airway triggered specific animations and responses. The environment and objects were then integrated with VR hardware, allowing participants to interact intuitively within the simulation. This comprehensive approach ensured a seamless blend of immersive VR elements and realistic interactions.

AI development tools were critical for implementing ML algorithms, and enhancing data analysis and interactivity in the VR environment. Visual Studio Code (Microsoft Corporation) served as the primary Integrated Development Environment (IDE) for developing and testing AI algorithms, with Pycharm (JetBrains) and Ubuntu (Canonical) used for debugging specific tasks. Python (Python Software Foundation) was chosen as the primary programming language due to its extensive library ecosystem, high portability, and flexibility, making it ideal for developing advanced ML projects. These AI tools enabled features such as speech recognition, semantic analysis, and real-time performance evaluation, allowing the system to capture and interpret participant interactions comprehensively.

By integrating these VR and AI development tools, the project successfully created a sophisticated and interactive training environment. This integration not only enhanced the realism of the VR scenarios but also provided robust

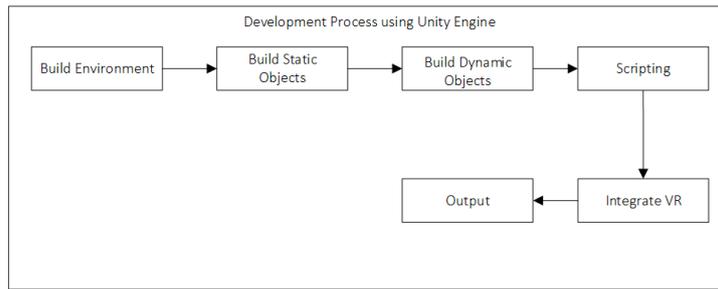


Figure 3. Unity Scripting Workflow

tools for analyzing participant performance, making it an effective platform for paramedic student training in Mass Casualty Incident scenarios.

Component 2: VR Learning Tool

The VR learning tool consists of a data collection platform designed to capture and analyze a wide range of participant interactions and performance metrics during VR training sessions. The platform integrates both hardware and software elements, including VR devices, media data collection tools, and a high-performance desktop computer. These components work together to ensure precise tracking, recording, and processing of data, providing a comprehensive foundation for evaluating the effectiveness of the VR learning tool.

Data Collection Platform

The data collection platform comprised hardware and software components designed to capture a wide range of data for evaluating participant performance and interaction with the VR learning tool. The hardware included motion sensors, a VR headset (HMD), media data collection devices, and a high-performance desktop computer, each playing a critical role in ensuring accurate data capture and analysis. Figure 4 shows the details of the training environment setup.

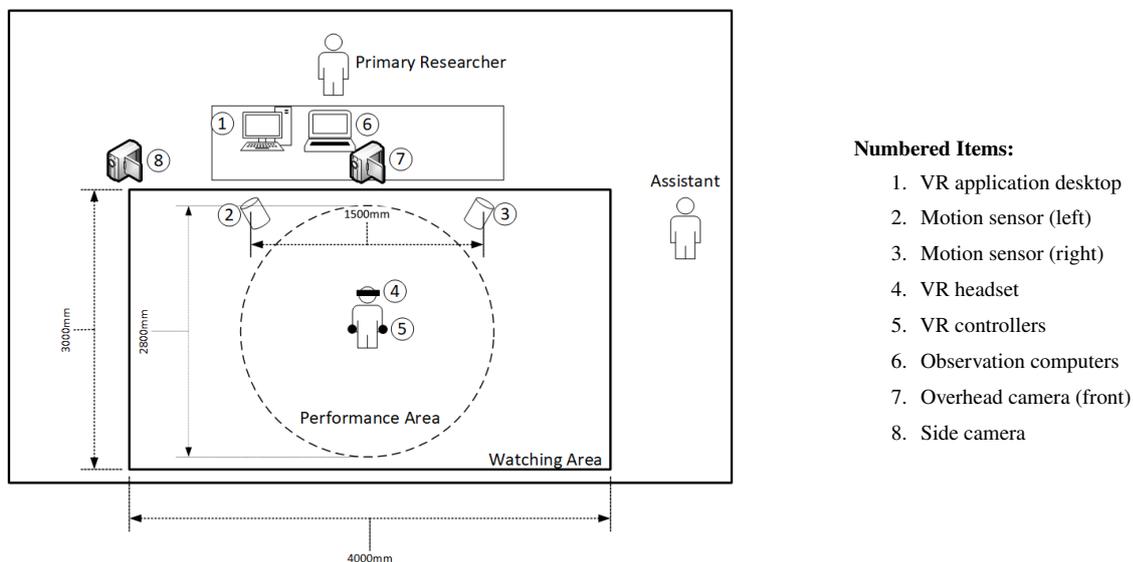


Figure 4. Training environment setup for the VR MCI simulation.

The primary VR device was the Oculus Rift VR headset (Item 4), featuring dual screens for stereoscopic 3D visuals. Embedded sensors within the headset tracked head movements, dynamically adjusting the visuals to create an immersive 3D environment for users. Supporting the headset were Oculus motion sensors (Item 2, 3), which enhanced motion tracking capabilities. Two motion sensors were deployed in this study to support room-scale VR training, ensuring precise monitoring of user actions. Additionally, the Oculus touch Controllers (Item 5) allowed participants to interact naturally within the virtual environment, enabling realistic and intuitive engagement with the training scenarios.

Media data collection devices were strategically positioned around the training area to capture participant interactions comprehensively. A wide-angle camera (Item 7) was mounted at a height of 2000mm at the front of the participants, while side cameras (Item 8) were positioned at chest height (1600mm) on either side. A voice recording device (Item 6) was placed directly in front of the participant to capture audio inputs during the sessions, such as verbal reports and interactions. The desktop computer (Item 1), configured to meet the rigorous demands of both VR system operation and software development, served as the backbone of the data collection platform. Equipped with an Intel Core i5 processor, 32GB of RAM, and an NVIDIA GeForce RTX 2070 Super graphics card, the desktop facilitated real-time processing, storage, and analysis of the collected data.

Data Evaluation Process

The evaluation process followed a structured workflow to assess the performance and effectiveness of the VR learning tool (Figure 5). This workflow consisted of two stages: transforming raw data into processed data, and subsequently deriving performance evaluation results through a structured analysis framework. An *Evaluation Matrix* was developed to systematically assess participants' performance across multiple dimensions. This matrix integrates both quantitative metrics—such as task completion time, casualty categorization accuracy, and scenario sequencing—and qualitative measures, including semantic coverage of METHANE reports and subjective feedback. It supports a comprehensive comparison across participants and scenarios, providing a basis for formative and summative evaluation.

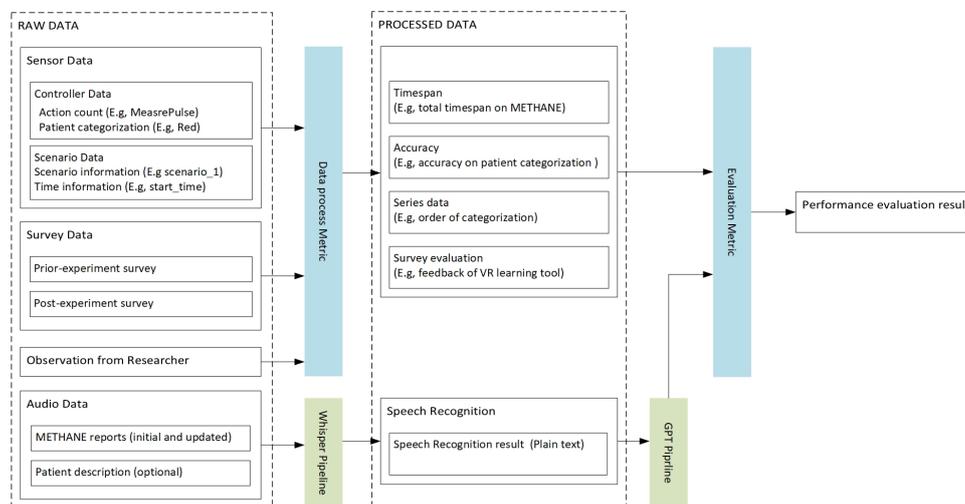


Figure 5. Data Evaluation Process Workflow

Raw data were obtained from four primary sources: sensor data, survey data, audio data, and researcher observations. Sensor data captured participant interactions facilitated by the Oculus touch controllers, including actions such as measuring casualties' vital signs and assigning triage categories. Scenario data, comprising timestamps and scenario indices, were collected to contextualize these interactions within the training framework. Survey data were gathered through pre-experiment and post-experiment surveys. The pre-experiment surveys collected participants' demographic information, prior exposure to VR, and familiarity with MCIs, while the post-experiment surveys solicited feedback on the effectiveness of the VR learning tool and suggestions for potential improvements.

Audio data constituted a significant dataset, emphasizing participants' verbal METHANE reports, which detailed incident characteristics, including the number and condition of casualties. These audio inputs were processed using the Whisper automatic speech recognition (ASR) model, which transformed the audio recordings into structured textual time-series data for further analysis. Subsequently, semantic analysis of these textual reports was performed utilizing the GPT language model, enabling the evaluation of their precision and completeness relative to predefined scenario descriptions. Researcher observations complemented these data sources, offering qualitative insights into participant behaviours, including their levels of engagement and communication styles, which suggested avenues for tailored training methodologies.

Processed data derived from these sources provided metrics such as task completion times, correctness in casualty categorization, and the sequencing of triage decisions. Statistical techniques were employed to analyze sensor, survey, and observational data, while advanced AI tools such as Whisper facilitated the comprehensive examination of audio inputs. The integration of quantitative metrics and qualitative feedback yielded a detailed evaluation of participant performance and the VR tool's overall training effectiveness.

The integration of hardware, software, and evaluation methodologies within the data collection platform directly addressed the study's objectives. It contributed to Research Question 2 by demonstrating the use of AI to enhance data analysis and supported Research Question 3 by assessing and improving the training effectiveness of the VR learning tool. The findings informed iterative refinements, ensuring the system's adaptability and utility in emergency healthcare training.

While subjective feedback remains a valuable component of training evaluation, this study advances beyond traditional approaches by adopting a multi-modal, data-driven evaluation framework. The evaluation process integrates objective sensor data (e.g., task completion time, participant interaction sequences) and quantitative analysis of structured speech content using Whisper and GPT-based semantic evaluation. By combining time-stamped interaction logs, speech, and researcher observations, this framework provides a more comprehensive evaluation of participant performance in complex MCI simulations.

Component 3: Training Experiment

The training experiment was designed to evaluate the performance of VR technology in enhancing triage training, with a focus on participant recruitment, the establishment of an optimized training environment, and a comprehensive approach to data collection.

Recruitment Process

The recruitment process is conducted to identify paramedic students (20–65 years, fluent in English, with practical emergency-care experience) who are suitable for triage training. Posters at Auckland University of Technology (AUT) campuses prompted interested individuals to contact the primary researcher. A three-week consideration period was offered with non-response interpreted as declining. Exclusion criteria (visual impairments, photosensitive epilepsy, or insufficient triage knowledge) are also established to protect the participant from potential illness and ensure the validity of the data collection. The recruitment process successfully enrolled 10 participants.

Training Sessions Setup

The car crash and earthquake scenarios were selected due to their relatively high frequency and different levels of complexity in MCIs. Car crashes are common high-pressure events that offer clear visual cues for triage, while earthquakes require broader situational awareness and present more ambiguous conditions, challenging participants to make complex decisions. Earthquakes are also frequent natural disasters in countries like New Zealand, making them highly relevant for local emergency response training.

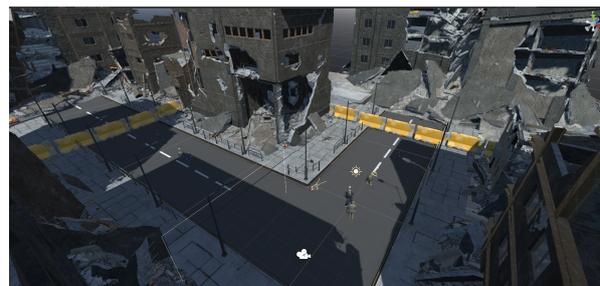
The sessions took place in a VR facility at AUT South Campus featuring a *performance area* and a surrounding *watching area* to ensure participant safety (See Figure 4).

Each training session comprised three scenarios:

- Scenario 0: a tutorial scenario to familiarize participants with the VR interface and operations.
- Scenario 1: a car crash scenario (Figure 6a) with 9 casualties and two METHANE report tasks.
- Scenario 2: an earthquake scenario (Figure 6b) with 6 casualties and two METHANE report tasks.



(a) Scenario 1: Car crash scenario



(b) Scenario 2: Earthquake scenario

Figure 6. Side-by-side comparison of Scenario 1 (Car crash) and Scenario 2 (Earthquake).

In scenarios 1 and 2, the participant was instructed to perform the initial METHANE report, triage the individual casualties and then complete the updated METHANE report. Each session lasted between 25–30 minutes, with scheduled breaks between scenarios. These breaks allowed participants to provide feedback on their experiences and report any discomfort. The principal researcher facilitated discussions during these intervals to gather qualitative insights and ensure the participant’s readiness to continue.

RESULTS

This section presents the findings of the study, offering a comprehensive evaluation of participant performance, user experiences, and communication effectiveness within the VR training environment. Among the 10 enrolled participants, 8 completed all training scenarios and 2 partially completed them. Data collected from the VR learning tool provided detailed insights into participants’ time management, correctness in triage tasks, and decision-making processes across various emergency scenarios. Survey results captured participants’ perceptions of the VR training, including its immersive quality, usability, and their overall acceptance of the technology. Additionally, advanced AI models were employed to analyze participants’ speech reports, assessing their adherence to the METHANE framework and identifying areas for improvement in structured communication. These results collectively highlight the strengths and challenges of using VR and AI technologies in emergency healthcare training and provide actionable insights for enhancing future implementations.

Sensor data results

The sensor data collected during the VR training sessions provided a comprehensive view of participant performance, focusing on time management, correctness in triage categorization, and the sequence of triage actions. Using embedded scripts within the VR learning tool, key controller actions were captured and consolidated into detailed datasets for evaluation. We use only the sensor data from the 8 participants who completed all scenarios to ensure completeness.

Scenario 1: Car Crash Accident

Analysis of the car crash scenario revealed notable variability in the time participants required to complete the triage tasks. There are 11 tasks in Scenario 1, including two METHANE reports and triage of nine individual casualties. The first task is delivering an initial METHANE report upon arrival, followed by triaging nine virtual patients, and concluding with an updated METHANE report once all assessments are completed. Correctness refers specifically to the accuracy of the nine triage decisions—measured by comparing each participant’s assigned priority level to a predefined expert classification. The two METHANE reports were assessed qualitatively for completeness but were not included in the numeric correctness percentage.

Table 1. Time span, Rankings, and Correctness of Participants in Scenario 1

Participant	SC1 Time span (s)	SC1 Time span rank	SC1 Correctness (%)
P2	487	3rd	66.7
P4	561	5th	33.3
P5	572	6th	77.8
P6	405	1st	66.7
P8	635	7th	33.3
P9	425	2nd	66.7
P10	459	4th	77.8
P11	757	8th	77.8
Average	537.6		62.5
Std. Dev.	118.5		18.7

Table 1 shows the total time spans, time span rank, and triage correctness for each participant, where Px refers to Participant x. The total time spans ranged from 405 seconds (for P6) to 757 seconds (for P11), with an average of 537.6 seconds. This average met the general triage time frame requirement, suggesting overall adherence to performance expectations. Triage correctness across participants showed a generally consistent trend, with most correctly identifying the level of six or more casualties. However, two participants reported pretty poor correctness, correlating with symptoms of nausea and dizziness experienced during the VR session. This finding highlights the impact of physical discomfort on cognitive performance and decision-making, emphasizing the importance of

Table 2. Time Spans for Individual Tasks in Scenario 1

Participant	Init. M	C1	C2	C3	C4	C5	C6	C7	C8	C9	Updt. M
P2	53	94	91	34	14	28	30	18	14	77	34
P4	30	85	82	34	65	51	34	19	29	19	113
P5	101	54	69	19	45	47	30	47	46	47	67
P6	69	62	49	17	19	33	28	30	8	17	73
P8	82	127	175	0	18	22	26	26	64	43	52
P9	83	40	33	40	35	18	15	16	43	35	67
P10	41	48	62	19	74	38	27	32	40	38	40
P11	0	148	149	89	70	24	35	15	33	38	156
Average	57.4	82.3	88.8	31.5	42.5	32.6	28.1	25.4	34.6	39.3	75.3
Std. Dev.	33.0	39.0	49.2	26.5	24.7	11.9	6.2	10.8	18.0	18.6	40.7

ergonomic and user-centered VR design. The Pearson correlation coefficient of time span and correctness is -0.13, indicating a weak negative correlation between time span and correctness.

Table 2 shows the time spans for each individual tasks. Here, Init. M refers to the initial METHANE report, Updt. M refers to the updated METHANE report, and Cx refers to Casualty x. The results indicate that C1 and C2 required notably more time than the other casualties and have higher standard deviations. This is due to two main reasons: (1) they were always firstly triaged; (2) the tasks associated with these two casualties were more complex. For example, Casualty 2 (C2) requires additional procedural steps in applying a tourniquet and performing airway assessments, indicating the complexity of this case. Due to different finish order and performance, the task completion time of the rest of the casualties show variety distribution. A closer examination of the (Initial and Updated) METHANE tasks revealed disparities in how participants approached the initial and updated assessments. Participants on average spent more time on the updated METHANE, but 5 out of 8 actually spent less time on the updated METHANE. The higher average is due to P4 and P11 spending significant time on the updated METHANE. So in general, we can see that there is less time spent on the updated METHANE than the initial one.

Scenario 2: Earthquake

The scenario 2 consists of eight predefined tasks, including two METHANE reports and the triage of six virtual casualties. In this earthquake scenario, participants exhibited greater consistency in time spans compared to Scenario 1. The total time spans ranged from 230 seconds for Participant 4 to 431 seconds for Participant 9, with an average of 299.5 seconds (See Table 3). This is the result of an improved familiarity with the VR environment and the less casualties in scenario 2. Moreover, the highest and lowest casualty triage times in Scenario 2 were both reduced compared to Scenario 1 (See Table 4). Participants displayed varied prioritization strategies, with Casualty 4 receiving the most attention due to its higher complexity. On average, participants allocated 47.3 seconds to Casualty 4, whereas Casualty 2 and 6 required much less time, suggesting their assessments were less demanding. Furthermore, while the initial METHANE reporting times remained similar between the two scenarios, the updated METHANE times in Scenario 2 tended to decrease, reflecting potentially more matured processes.

Table 3. Time span, Rankings, and Correctness of Participants in Scenario 2

Participant	SC2 Time span (s)	SC2 Time span rank	SC2 Correctness (%)
P2	314	4th	50.0
P4	230	1st	50.0
P5	289	3rd	50.0
P6	271	2nd	33.3
P8	264	5th	33.3
P9	431	8th	50.0
P10	322	6th	50.0
P11	275	7th	33.3
Average	299.5		43.7
Std. Dev.	60.5		8.6

Triage correctness in Scenario 2 was lower than in Scenario 1, with most participants correctly identifying only three casualties out of six (See Table 3). The findings suggest a need for enhanced training and support to improve decision-making in more challenging scenarios.

Table 4. Time Spans of Individual Tasks in Scenario 2

Participant	Init. M	C1	C2	C3	C4	C5	C6	Updt. M
P2	71	25	21	9	43	32	32	81
P4	45	18	19	14	42	33	25	34
P5	51	30	22	17	60	28	16	65
P6	47	18	13	30	27	40	25	71
P8	62	21	12	22	20	48	24	55
P9	94	103	22	60	27	19	19	87
P10	57	60	23	32	54	31	23	42
P11	31	39	26	16	105	4	2	52
Average	57.3	39.3	19.8	25.0	47.3	29.4	20.8	60.9
Std. Dev.	19.1	29.3	4.9	16.2	27.1	13.3	8.9	18.5

Comparative analysis across the two scenarios revealed patterns of consistency and adaptability among participants. Some, like Participant 6, ranked 1st and 2nd in time span across both scenarios, demonstrating consistent efficiency (Table 4). Others, such as Participant 4, shows a significant improvement in time span ranking in Scenario 2, suggesting that practice or adaptation to specific scenario requirements can enhance performance. On the other hand, participants like Participant 9 experienced a notable decline of time span ranking, highlighting challenges in adapting to differing scenario demands.

Overall, the sensor data indicates the importance of balancing efficiency with thoroughness in triage tasks. It also highlighted the critical role of physical comfort in a VR environment, as discomfort can negatively impact decision-making and performance. These findings provide valuable insights for refining VR-based training programs, emphasizing the need for user-friendly interfaces and tailored training to optimize outcomes in emergency healthcare scenarios.

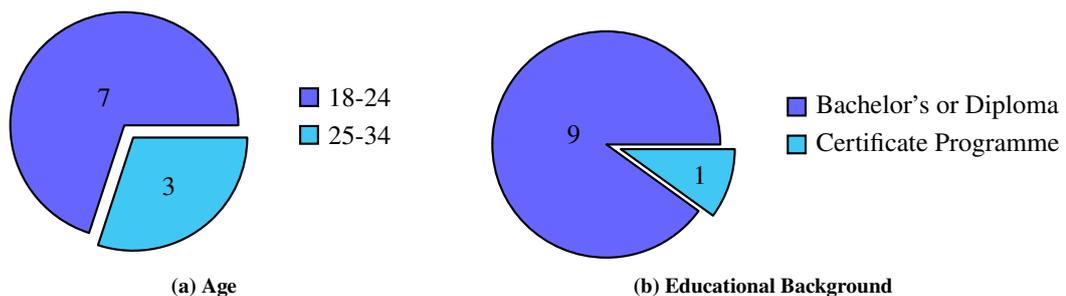
While basic time tracking and task accuracy are common in training studies, this research extends prior work by capturing detailed interaction data within a VR environment. By linking time-stamped actions to triage decisions and verbal reports, the system provides deeper insight into participant decision-making flow—going beyond traditional metrics to support more comprehensive performance analysis.

Survey Results

The survey results provided valuable insights into the participants' backgrounds, their experiences during the VR training sessions, and their perceptions of the training's effectiveness. Both the initial and post-experiment surveys were carefully designed to evaluate participants' prior experiences and gather detailed feedback on their interactions with the VR learning tool.

Initial Survey

The initial survey, comprising four questions, collected essential demographic and experiential data. The participants, predominantly aged 18–24 (seven participants) with the remainder aged 25–34 (three participants), represented a younger demographic comfortable with digital tools (Figure 7a). In terms of educational background, nine participants held a bachelor's or undergraduate diploma, while one participant had completed a certificate programme (Figure 7b).

**Figure 7. Survey Results: Age and Educational Background Distributions**

The participants’ familiarity with VR technology varied widely (Figure 8a): three were new users with no prior VR experience, six had some experience with fewer than ten uses, and one was an experienced user with more than ten uses. This diverse range of familiarity indicates the challenge of accommodating varying experience levels within the group.

Participants’ prior knowledge of MCI training methods varied (Figure 8b). All participants gained theoretical knowledge from lectures and textbooks, six participants engaged in paper-based simulations, and eight participants participated in live simulations such as AUT MI Day. However, none had experience with VR-based simulation.

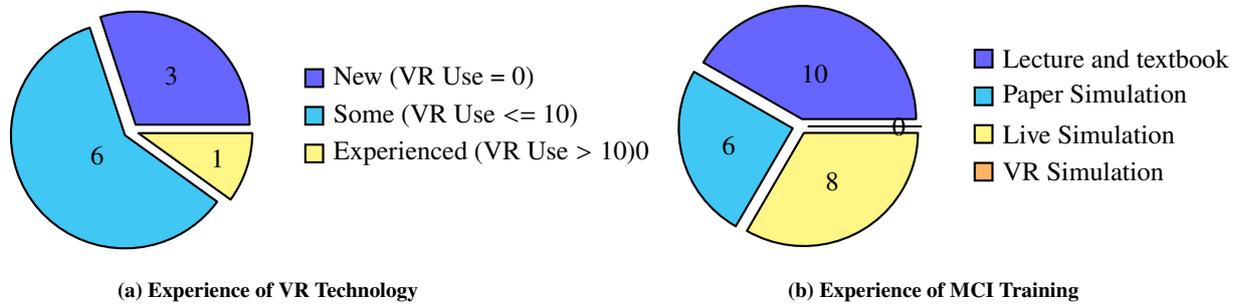


Figure 8. Survey Results: VR Technology and MCI Training Experiences

Post-Experiment Survey

The post-experiment survey included 16 questions designed to evaluate participants’ immersion, engagement, usability, skill confidence, physical well-being, and the overall effectiveness of the VR learning tool. Among those 16 questions, 12 used a Likert scale to measure responses, while 4 were open-ended, allowing detailed answers. These questions aimed to capture participants’ experiences and gather feedback for potential improvements. The results for Q1-Q12 are summarized in Tables 5. We assigned a score of 5 to “Strongly Agree” at the top of the scale and a score of 1 to “Strongly Disagree” at the bottom.

Table 5. Survey Feedback Summary

Question	Strongly Agree (5)	Agree	Neutral	Disagree	Strongly Disagree (1)	Average	Std. Dev.
Q1 (Immersion)	5	4	1	0	0	4.4	0.66
Q2 (Presence)	3	4	3	0	0	4.0	0.77
Q3 (Presence)	3	6	0	1	0	4.1	0.83
Q4 (Engagement)	6	4	0	0	0	4.6	0.49
Q5 (Emotion)	2	5	3	0	0	3.9	0.70
Q6 (Usability)	1	4	5	0	0	3.6	0.66
Q7 (Skills)	4	4	2	0	0	4.2	0.75
Q8 (Skills)	4	4	1	1	0	4.1	0.94
Q9 (Nausea)	2	1	1	3	3	2.6	1.50
Q10 (Dizziness)	0	2	2	3	3	2.3	1.10
Q11 (Adoption)	10	0	0	0	0	5.0	0.00
Q12 (Adoption)	10	0	0	0	0	5.0	0.00

• Immersion and Presence (Q1-Q3)

In Q1, nine participants (five strongly) agreed that the VR learning tool was more immersive than paper-based MCI triage simulations. In Q2, seven participants (three strongly) agreed that the VR environment felt natural to interact with. In Q3, nine participants (three strongly) agreed that they were able to examine the objects closely in the VR environment. These results suggest that participants generally found the VR learning tool to provide an immersive, realistic, and interactive experience, particularly regarding natural interaction and close object examination.

• Engagement and Usability (Q4-Q6)

In Q4, all ten participants (six strongly) agreed that they felt engaged in the virtual environment experience. In Q5, seven participants (two strongly) agreed that they felt tense in the VR environment, while three participants were neutral. In Q6, five participants (one strongly) agreed that the interaction devices were easy to use, and the remaining five participants were neutral. The results suggest that the VR environment effectively engaged participants, while there is room for improvement in tension reduction and the usability of interaction devices.

- Skill Confidence (Q7-Q8)

In Q7, eight participants (four strongly) agreed that they felt confident selecting objects in the VR environment, while two participants were neutral. In Q8, eight participants (four strongly) agreed that they were confident using the controller to move around, one participant was neutral, and one disagreed. The results suggest that most participants felt confident in both selecting objects and using the controller to move around in the VR environment.

- Physical Well-Being (Q9-Q10)

In Q9, three participants (two strongly) agreed that they suffered from nausea, one participant was neutral, and six participants disagreed (three strongly). In Q10, two participants agreed they suffered from dizziness, two remained neutral, and six disagreed (three strongly). The result shows that only a minority of participants experienced nausea or dizziness, while the majority felt neutral or better.

- Technology Adoption (Q11-Q12)

In Q11, all ten participants strongly agreed that the VR system can be an integral tool for MCI training. In Q12, likewise, all ten participants strongly agreed that they would like to see more VR-based MCI scenarios, such as different incidents or disasters. The results indicate that all participants strongly recommend the value of the VR learning tool for MCI training.

- Feedback and Suggestions

Open-ended questions revealed areas for improvement, including slower movement speeds to reduce nausea and more realistic physics for object interaction. Participants also suggested dynamic elements, such as interactive casualties and environmental sound effects, to enhance realism. Additionally, the inclusion of wireless headsets and larger physical spaces was recommended to improve usability and safety during training.

For Questions 1–8 and 11–12, a higher mean (out of 5) corresponds to a more positive evaluation (e.g., greater immersion, usability, or skill confidence). In contrast, for Questions 9 (Nausea) and 10 (Dizziness), a higher mean represents more reported discomfort. The standard deviation (SD) shows how much individual responses vary: Q9 and Q10 have higher SD values, indicating greater variability in perceived physical comfort, while lower SD values in the other questions suggest more uniform responses.

Overall, the survey results demonstrate a generally positive reception to the VR learning tool, highlighting its immersive and engaging nature. Nonetheless, issues such as physical discomfort and a steeper learning curve for some users point to areas for further refinement. Implementing improvements in motion mechanics, realism, and interactivity could significantly enhance the effectiveness of VR training within emergency healthcare scenarios.

Speech Recognition and Semantic Similarity Evaluation

This study employed advanced AI models to evaluate participants' verbalized reports during a VR-based MCI training scenario, focusing on adherence to the METHANE framework. The OpenAI Whisper algorithm was used to transcribe participants' speech data. These transcriptions provided a rich dataset for analysis, exemplified by Table 6, which displays sample excerpts of participant speech during the simulation.

Table 6. Example of speech content based on time series data.

TimeSeries	Speech content
01:10:22 PM	If I start my METHANE report now, I say major incident declared.
01:10:27 PM	Exact location is the, just before the Papakura exit 458 on State Highway 1.
01:10:34 PM	Type of incident is an MVA, motor vehicle accident.
01:10:38 PM	Hazards include traffic in both lanes on either side.

METHANE evaluation using GPT model

The OpenAI GPT4 model was employed to assess how comprehensively participants covered each METHANE component in their reports. The GPT4 model uses an expert METHANE report (see Table 7) for Scenario 1, as the ground truth reference for assessing component coverage in participants' METHANE reports. The initial and updated METHANE reports provided by participants were analysed and categorised as Full, Part, or None for each METHANE component. The initial part of the GPT prompt is presented in Figure 9. This categorization was based on several pre-defined rules, which compared the participant's speech transcripts to the predefined expert report for

Paramedic Expert Judgment Task

Suppose you are a paramedic expert, and you will judge the METHANE reports from a participant in a VR experiment on MCI triage.

There are two reports: the **initial METHANE report** at the beginning of the experiment and the **updated METHANE report** at the end of the experiment.

You will be given two sample reports by an expert as the **ground truth**, and you will make your judgment according to these reports.

In each judgment, you will report whether each **M E T H A N E** element is:

- **Fully covered** (Full)
- **Partially covered** (Partial)
- **Missing** (None)

Figure 9. GPT Evaluation Prompt for Judging METHANE Reports

each scenario. Each METHANE component was rated as “Full” if it was explicitly and accurately covered, “Partial” if it was mentioned but lacked detail or precision, and “None” if it was not mentioned entirely.

This approach enabled a more automated and scalable analysis of communication quality than traditional free-text scoring methods. The results are shown in Table 8.

Table 7. The expert METHANE report for Scenario 1

	Initial METHANE report	Updated METHANE report
M	Ambulance to control, I would like to declare a major incident.	Ambulance to control, I would like to confirm a major incident.
E	The location of the incident is on State Highway 1 towards junction 458, heading southbound.	The location of the incident remains on State Highway 1 towards junction 458, heading southbound.
T	The incident is a multi-vehicle Accident.	The incident is a multi-vehicle accident involving 3 number of vehicles.
H	There is debris like glass on the road.	There is debris like glass on the road.
A	Access is via a southbound entry point before Junction 458, with egress at Junction 458.	Access is via a southbound entry point before Junction 458, with egress at Junction 458.
N	There is an estimated nine casualties.	There is an estimated nine casualties. We have the following casualties’ status codes: 2 status zero casualties, 2 status Ones casualties, 3 status Twos casualties, 2 status Three casualties.
E	There are currently no other emergency services on scene. Please send fire to contain hazards and Police to control traffic and eight further ambulances.	Police and fire have not yet arrived. We require an addition 2 number of ambulances and please send Westpac HEMs (helicopter).

Initial METHANE Reports:

The analysis of initial reports revealed considerable variability among participants in both coverage and detail. Participant 6 demonstrated exceptional performance, achieving full coverage across all METHANE components. Their report included precise declarations of the major incident, location details, incident type, hazards, access routes, casualty numbers, and emergency service needs. This consistency indicated a strong understanding of the METHANE communication protocol and its application.

In contrast, Participant 4’s initial report lacked coverage across most components, omitting critical details such as the specific hazards, access points, and required emergency services. This significant gap highlighted challenges in effectively communicating essential information during the initial phase of the scenario. Similarly, Participant

11 failed to submit an initial report, a substantial omission given the importance of structured communication in emergency responses.

Partial coverage was observed in the reports of other participants. For example, Participant 2 addressed components like Access and emergency services, but their descriptions lacked clarity and completeness, failing to meet the standard set by the predefined METHANE report. Participant 8 also provided partial coverage, with vague descriptions of the incident type and casualty numbers, indicating room for improvement in articulating critical details.

Updated METHANE Reports:

A notable improvement was observed in the updated reports, with most participants addressing previously missed components and providing greater detail. Participant 4, who had significant gaps in their initial report, achieved full coverage in their updated submission. This improvement demonstrated an increased understanding of the METHANE framework and its application in the scenario. Participant 2 also enhanced their report, addressing the Access and Casualty Numbers components more comprehensively, although some minor gaps remained.

Participants such as 10 and 6 maintained high performance in their updated reports, providing consistent and detailed coverage across all METHANE components. Participant 6’s updated report reiterated key details from the initial submission while adding more specific information, such as the breakdown of casualty statuses, aligning closely with the predefined METHANE report.

Participant 11, despite missing their initial report, delivered a partially complete updated report. While they addressed components like Major Incident Declaration and Casualty Numbers, their descriptions of Location and Incident Type lacked precision. This indicates the ongoing need for practice and feedback to strengthen reporting skills.

Table 8. Individual METHANE component results for Scenario 1

Participant	Type	M	E	T	H	A	N	E
P2	Initial	Partial	Full	Full	Full	Partial	Full	Partial
	Updated	Full	Full	Full	Full	Partial	Full	Full
P4	Initial	Partial						
	Updated	Full						
P5	Initial	Partial	Full	Full	Full	Partial	Full	Full
	Updated	Full						
P6	Initial	Full						
	Updated	Full						
P8	Initial	Partial	Partial	Partial	Partial	Partial	Full	Full
	Updated	Partial	Full	Full	Full	Full	Partial	Full
P9	Initial	Partial	Partial	Full	Full	Full	Full	Full
	Updated	Full	Full	Full	Full	Full	Full	Partial
P10	Initial	Full	Full	Full	Full	Partial	Partial	Full
	Updated	Full						
P11	Initial	Partial						
	Updated	Full	Partial	Partial	Full	Partial	Partial	Full

Legend: Full coverage Partial coverage No coverage

Summary and validation

Table 8 highlights the progression of participants in covering METHANE components. Many of the initial reports lacked completeness, but the updated METHANE reports demonstrated notable improvements in both coverage and detail following iterative practices. Those who adhered closely to the METHANE structure produced clearer, more actionable reports, illustrating the value of standardized communication protocols in emergency scenarios. To further enhance VR-based training, efforts should focus on commonly missed elements (such as hazards and emergency services), provide detailed evaluations after each METHANE report, and incorporate practice under simulated stress conditions (e.g., time constraints and environmental distractions).

Unlike previous studies that relied solely on manual scoring or speech transcription, this study employed GPT-4 for semantic evaluation, enabling a more automated and scalable analysis of communication quality. GPT-4 was used to assess the completeness and accuracy of METHANE reports by comparing them to predefined expert scenario descriptions, offering a structured approach to quantify adherence to reporting protocols. To validate the reliability of GPT-4's assessments, a paramedic expert independently reviewed and validated the METHANE evaluations alongside the research team. The validation results indicated a high level of consistency between GPT-4 and the expert, with strong agreement across most METHANE components. Additionally, although the current sample size (N=8) was sufficient for a preliminary investigation, expanding the cohort in future research would help to reinforce the robustness and generalisability of the conclusions.

CONCLUSION AND FURTHER RESEARCH

This study investigated the integration of VR and AI technologies in pre-hospital triage training for MCIs, demonstrating their potential to improve emergency healthcare education. The findings highlighted how immersive VR environments, paired with AI-driven performance assessments, provide a robust platform to evaluate participant competency. Quantitative analysis, such as on task completion time and qualitative analysis such as on participant feedback offered a deeper understanding of user behaviours and decision-making under in the MCIs simulation and training. The use of advanced VR controllers facilitated realistic interactions, enabling behaviours that closely mirrored those required in real-world emergencies. This immersive quality proved instrumental in accurately assessing participant readiness and decision-making capabilities. Although participants generally accepted VR technology positively, initial unfamiliarity challenges indicate the importance of comprehensive onboarding and support to maximize engagement and learning outcomes.

However, the study recruited 10 participants, providing initial insights but limiting statistical depth and representativeness. While the VR scenarios offered a useful foundation for examining real-world-like behaviours, expanding scenario diversity could improve behavioural analysis. Although creating realistic vital sign models and quantifying triage within VR posed challenges, iterative refinements helped approximate real-world conditions.

Overall, this study shows the potential of VR and AI in emergency healthcare training. By addressing existing limitations and advancing the technological capabilities of VR learning tools, future research recommends an inclusive recruitment strategy to increase participant diversity and deepen the study's findings. Expanding both the sample size and the complexity of training scenarios would offer a more comprehensive view of how VR influences participant behaviour across various emergencies. Advancing vital sign modelling—potentially through dynamic simulations and ML—and refining casualty interactions would boost realism and educational impact. Moreover, incorporating image-based pattern recognition methods (e.g., video data evaluation and pose recognition) would enable a more comprehensive assessment of participant performance. Finally, diversifying scenarios (such as natural disasters or industrial accidents) would further enhance the VR tool's applicability to real-world conditions and help healthcare professionals prepare for a broader range of emergency situations.

REFERENCES

- Aggarwal, R., Grantcharov, T. P., Eriksen, J. R., Blirup, D., Kristiansen, V. B., Funch-Jensen, P., & Darzi, A. (2006). An evidence-based virtual reality training program for novice laparoscopic surgeons. *Annals of surgery*, 244(2), 310–314.
- Andreatta, P., Maslowski, E., Petty, S., Shim, W., Marsh, M., Hall, T., Stern, S., & Frankel, J. (2010). Virtual reality triage training provides a viable solution for disaster-preparedness. *Academic Emergency Medicine*, 17(8), 870–876. <https://doi.org/10.1111/j.1553-2712.2010.00728.x>
- Behmadi, S., Asadi, F., Okhovati, M., & Sarabi, R. E. (2022). Virtual reality-based medical education versus lecture-based method in teaching start triage lessons in emergency medical students: Virtual reality in medical education. *Journal of Advances in Medical Education & Professionalism*, 10(1), 48.
- Berndt, H., Wessel, D., Mentler, T., & Herczeg, M. (2018). Human-centered design of a virtual reality training simulation for mass casualty incidents. *Proceedings of the 2018 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. <https://doi.org/10.1109/VS-Games.2018.8493427>
- Berndt, H., Wessel, D., Willer, L., Herczeg, M., & Mentler, T. (2018). Immersion and presence in virtual reality training for mass casualty incidents. *ISCRAM*.
- Billis, A., Papadimitriou, G., Douligeris, C., Papavramidis, T., Krizea, M., Bamidis, P., Logaras, E., Zouka, M., Karanasiou, N., Fourlis, A., Nicopolitidis, P., Lagakis, P., Gialelis, J., & Kallergis, D. Functional and non-functional requirements of a smart triage system for emergency departments: The case of inteltrriage project. In: 2019. <https://doi.org/10.1109/seedac-censm.2019.8908320>

- Cabrera, E., Luque, E., Taboada, M., Epelde, F., & Iglesias, M. Optimization of emergency departments by agent-based modeling and simulation. In: 2012, 423–430. <https://doi.org/10.1109/iri.2012.6303040>
- Caicedo-Torres, W., García, G., & Pinzón, H. (2016). A machine learning model for triage in lean pediatric emergency departments. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10022 Lnai, 212–221. https://doi.org/10.1007/978-3-319-47955-2_18
- Canonical. (2021, May 11). *Ubuntu* (Version 20.04).
- Cone, D., Serra, J., & Kurland, L. (2011). Comparison of the salt and smart triage systems using a virtual reality simulator with paramedic students. *European Journal of Emergency Medicine*, 18(6), 314–321. <https://doi.org/10.1097/MEJ.0b013e328345d6fd>
- Fan, X. (2012). Decision as choice of potential intentions. *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology - Volume 02*, 220–227. <https://doi.org/10.1109/wi-iat.2012.44>
- Foronda, C., Shubeck, K., Swoboda, S., Hudson, K., Budhathoki, C., Sullivan, N., & Hu, X. (2016). Impact of virtual simulation to teach concepts of disaster triage. *Clinical Simulation in Nursing*, 12(4), 137–144. <https://doi.org/10.1016/j.ecns.2016.02.004>
- Gaba, D. M. (2004). The future vision of simulation in health care. *Quality and safety in Health care*, 13(suppl 1), i2–i10.
- Guo, Y., Wei, J., Ademu, L., & Yang, F. (2024). Past, present, and future of virtual reality in healthcare education and training: Bibliometric analysis. *Journal of Public Health Issues and Practices*. <https://doi.org/10.33790/jphip1100227>
- Harada, S., Suga, R., Suzuki, K., Kitano, S., Fujimoto, K., Narikawa, K., Nakazawa, M., & Ogawa, S. (2024). Usefulness of self-selected scenarios for simple triage and rapid treatment method using virtual reality. *Journal of Nippon Medical School*, 91(1), 99–107.
- Ingrassia, P., Ragazzoni, L., Carengo, L., Colombo, D., Gallardo, A., & Corte, F. (2015). Virtual reality and live simulation: A comparison between two simulation tools for assessing mass casualty triage skills. *European Journal of Emergency Medicine*, 22(2), 121–127. <https://doi.org/10.1097/mej?0000000000000132>
- Izard, S. G., & Méndez, J. A. J. (2016). Virtual reality medical training system. *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality*, 479–485. <https://doi.org/10.1145/3012430.3012560>
- JetBrains. (2021, June 21). *PyCharm* (Version 2022).
- Kang, D.-Y., Cho, K.-J., Kwon, O., Kwon, J.-M., Jeon, K.-H., Park, H., Lee, Y., Park, J., & Oh, B.-H. (2020). Artificial intelligence algorithm to predict the need for critical care in prehospital emergency medical services. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, 28(1). <https://doi.org/10.1186/s13049-020-0713-4>
- Kerr, E., McGinnity, T., Coleman, S., & Shepherd, A. (2021). Human vital sign determination using tactile sensing and fuzzy triage system. *Expert Systems with Applications*, 175. <https://doi.org/10.1016/j.eswa.2021.114781>
- Kilmon, C. A., Brown, L., Ghosh, S., & Mikitiuk, A. (2010). Immersive virtual reality simulations in nursing education. *Nursing Education Perspectives*, 31(5), 314–7.
- Koutitas, G., Smith, K. S., Lawrence, G., Metsis, V., Stamper, C., Trahan, M., & Lehr, T. (2019). A virtual and augmented reality platform for the training of first responders of the ambulance bus. *Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, 299–302. <https://doi.org/10.1145/3316782.3321542>
- Larsen, C. R., Soerensen, J. L., Grantcharov, T. P., Dalsgaard, T., Schouenborg, L., Ottosen, C., Schroeder, T. V., & Ottesen, B. S. (2009). Effect of virtual reality training on laparoscopic surgery: Randomised controlled trial. *Bmj*, 338, b1802.
- Lineberry, M., Walwanis, M., & Reni, J. (2013). Comparative research on training simulators in emergency medicine: A methodological review. *Simul. healthc.*, 8(4), 253–261. <https://doi.org/10.1097/SIH.0b013e31828715b1>
- Maclure, J. (2021). AI, explainability and public reason: The argument from the limitations of the human mind. *Minds and Machines*, 31(3), 421–438. <https://doi.org/10.1007/s11023-021-09570-x>
- Mantovani, F., Castelnuovo, G., Gaggioli, A., & Riva, G. (2003). Virtual reality training for health-care professionals. *CyberPsychology & Behavior*, 6(4), 389–395.

- Marx, J., Hockberger, R., & Walls, R. (2013). *Rosen's emergency medicine-concepts and clinical practice e-book: 2-volume set*. Elsevier Health Sciences.
- Microsoft Corporation. (2021, May 18). *Visual Studio Code* (Version 1.91).
- Mills, B., Dykstra, P., Hansen, S., Miles, A., Rankin, T., Hopper, L., Brook, L., & Bartlett, D. (2020). Virtual reality triage training can provide comparable simulation efficacy for paramedicine students compared to live simulation-based scenarios. *Prehospital Emergency Care*, 24(4), 525–536. <https://doi.org/10.1080/10903127.2019.1676345>
- Mistry, D., Brock, C. A., & Lindsey, T. (2023). The present and future of virtual reality in medical education: A narrative review. *Cureus*, 15(12), e51124. <https://doi.org/10.7759/cureus.51124>
- Paslı, S., Şahin, A. S., Beşer, M. F., Topçuoğlu, H., Yadigaroglu, M., & İmamoğlu, M. (2024). Assessing the precision of artificial intelligence in emergency department triage decisions: Insights from a study with chatgpt. *The American Journal of Emergency Medicine*, 78, 170–175.
- Python Software Foundation. (2022, June 15). *Python* (Version 3.11).
- Sung, H., Kim, M., Park, J., Shin, N., & Han, Y. (2024). Effectiveness of virtual reality in healthcare education: Systematic review and meta-analysis. *Sustainability*, 16(19). <https://doi.org/10.3390/su16198520>
- Tahayori, B., Chini-Foroush, N., & Akhlaghi, H. (2021). Advanced natural language processing technique to predict patient disposition based on emergency triage notes. *EMA - Emergency Medicine Australasia*, 33(3), 480–484. <https://doi.org/10.1111/1742-6723.13656>
- Yan, S., Peng, J., Grain, H., & Yi, M. (2019). Technology road mapping of two machine learning methods for triaging emergency department patients in australia. *Proceedings of the 2019 the International Conference on Pattern Recognition and Artificial Intelligence*, 60–67. <https://doi.org/10.1145/3357777.3357779>