Journal of Digital Life 2026, 6, S1

Special Issue: Influence of the Metaverse on Digital Life
Article

Development of a VR Vertical Evacuation Simulator for Schools Integrating Conformity Bias for Disaster Preparedness Education

Masaya Nakahara [1*], Yoshinori Tsukada [2], Yoshimasa Umehara [3], Yutaka Tada [4], Yasunori Kozuki [5], Yukari Ando [6], Takumu Kuhara [7] and Marina Kobayashi [7]

[1]Faculty of Information Science and Arts, Osaka Electro-Communication University, 1130-70 Kiyotaki, Shijonawate-shi, Osaka 575-0063, Japan
[2]Faculty of Engineering, Reitaku University, 2-2-1 Hikarigaoka, Kashiwa-shi, Chiba 277-8686, Japan
[3]Faculty of Business Administration, Setsunan University, 17-8 Ikedanakamachi, Neyagawa-shi, Osaka 572-8508, Japan
[4]Graduate School of Science and Engineering, Ehime University, 3 BunKyo-cho, Matsuyama, Ehime 790-8577, Japan
[5]Center for Environmental and Disaster Resilience, Tokushima University, 2-1 Minamijosanjima, Tokushima, Tokushima 770-8506, Japan
[6]Graduate School of Disaster Resilience and Governance, University of Hyogo, The Great Hanshin-Awaji Earthquake Memorial Disaster Reduction and Human Renovation Institution, 1-5-2 Wakahama Kaigan-dori, Chuo-ku, Kobe, Hyogo 651-0073, Japan
[7]Graduate School of Information Science and Arts, Osaka Electro-Communication University, 1130-70 Kiyotaki, Shijonawate-shi, Osaka 575-0063, Japan

Abstract
In recent years, Japan has experienced frequent water-related disasters, including earthquake-induced tsunamis and flooding caused by heavy rainfall, increasing the importance of disaster preparedness measures. However, participation in disaster drills remains low, indicating that many citizens do not regard disaster preparedness as a personally relevant issue. In addition, conventional drills often require participants to follow predetermined routes and therefore make it difficult to experience unexpected situations, such as blocked routes or collective behavior during panic. Although disaster simulators using virtual reality (VR) have advanced, few studies have focused on indoor flooding while also considering the influence of other evacuees' incorrect evacuation behavior. This study therefore developed a simulator that enables users to experience vertical evacuation during indoor flooding in school facilities, which are often designated as evacuation sites. By placing non-player characters (NPCs) that exhibit incorrect evacuation behavior, the simulator makes it possible to examine how the behavior of others may affect participants' evacuation decisions, subjective difficulty, and sense of urgency. An empirical evaluation clarified the simulator's potential value for disaster preparedness education and awareness raising, as well as issues requiring further verification.

Keywords: Indoor Disaster Prevention Education, Flood Simulation, Evacuation Simulation, Conformity Bias, VR

1. Introduction

Japan is an island nation surrounded by the sea and has many rivers; consequently, it is frequently affected by water-related disasters, including tsunamis caused by earthquakes and flooding caused by localized torrential rainfall. During the January 2024 Noto Peninsula earthquake, tsunami waves and river flooding after the earthquake inundated approximately 190 ha, with a maximum depth of about 4 m (Ministry of Land, Infrastructure, Transport and Tourism, 2024). Reports of this disaster were disseminated not only throughout Japan but also internationally. In addition, public information campaigns regarding the Nankai Trough earthquake, which is considered highly likely to occur in

Received: 21 November 2025, Revised: 24 May 2026, Accepted: 22 July 2026, Published: 14 August 2026
* Correspondence: nakahara@osakac.ac.jp

the future, are repeatedly conducted (Japan Meteorological Agency, 2025). These circumstances have made disaster prevention and mitigation measures for protecting human life more important than ever. However, according to a nationwide survey on disaster prevention and mitigation awareness conducted in fiscal year 2024 (National Federation of Workers and Consumers Kyosai Cooperatives, 2024), only 4.4% of respondents reported that they “almost always participate” in disaster drills held in their local communities. Across all age groups from people in their 20s to those in their 60s, the largest share of respondents reported that they had “never participated” in such drills. These results suggest that, despite Japan’s repeated disaster experiences and the broad dissemination of disaster-related information, many citizens do not regard disasters as personally relevant threats. If this situation continues, people may fail to take appropriate action when a disaster occurs, increasing the likelihood of severe damage. Moreover, even when people participate in drills, many conventional evacuation drills focus on orderly movement along designated evacuation routes. They therefore do not sufficiently account for situations that are difficult to reproduce or experience in ordinary times, such as evacuation routes becoming blocked by collapsed structures or groups making incorrect decisions and choosing wrong routes because of panic during evacuation.

Against this backdrop, previous studies have developed simulators that reproduce disaster-specific phenomena together with disaster situations in Virtual Reality (VR) environments. By enabling participants to experience these situations, such simulators support training that assumes actual disaster conditions, and their effectiveness for disaster preparedness education has been evaluated. In outdoor settings in particular, simulators reproducing diverse disasters and situations have been proposed. In indoor settings, however, many simulators assume fires, terrorism, or earthquakes, whereas relatively few studies have addressed flooding caused by tsunamis or similar events. Furthermore, even simulators that target flooding rarely consider the influence of evacuees other than the participant. Therefore, this study targets schools, which are often designated as evacuation sites, and develops a simulator in which participants evacuate vertically from a flooded indoor area to the rooftop. The study investigates whether incorporating the behavioral patterns of NPCs other than the participant, and examining their influence on participants’ evacuation behavior and psychological state, can contribute to enhancing disaster preparedness awareness.

2. Methods

This chapter investigates existing research and initiatives on disaster prevention education and evacuation experiences using VR. It identifies issues that arise when such approaches are used for education aimed at improving disaster preparedness awareness and organizes possible responses to those issues. Based on these responses, we propose a VR disaster preparedness learning simulator intended to further enhance participants’ disaster preparedness awareness.

2.1 Existing Methods

Many simulators have been developed for disaster preparedness education and evacuation experiences, regardless of whether they target indoor or outdoor settings or specific disaster types. Outdoor simulators include a board-game-style simulator that reproduces fires, heavy rainfall, and various other disasters across an entire urban area (Shimbo et al., 2022); simulators for tsunami evacuation in congested urban streets and around schools (Arakawa et al., 2025; Marui et al., 2024); and a simulator that estimates evacuation delays from the perspective of people requiring assistance (Nakai et al., 2018). In the board-game-style simulator, disasters occur according to the space on which the player stops, making it possible to visualize disaster conditions at each location. This design enables even children to experience disaster situations while enjoying the format of a board game. The tsunami evacuation simulators targeting areas such as city centers and school surroundings allow users to evacuate from tsunami inundation while avoiding congestion in fields that reproduce real-world structures. These simulators help users learn evacuation methods and routes that are difficult to grasp during normal times. The evacuation-delay simulator for people requiring assistance enables users to confirm, on a Geographic Information System (GIS), the evacuation routes and approximate evacuation times needed by people such as home-care recipients who depend on medical devices. These examples show that outdoor simulators can reproduce a wide range of disasters and conditions.

Indoor simulators include evacuation-training simulators for medical personnel in hospitals (Farra et al., 2018; Kaggwa et al., 2025), as well as evacuation simulators for school fires (Lorusso et al., 2022), earthquakes (Mitsuhara, 2024), and flooding (Iinuma et al., 2024). Hospital evacuation-training simulators allow users to evacuate through a hospital in which newborns and patients requiring assistance are present, while also supporting those evacuees. This makes it possible to learn not only evacuation routes but also the actions necessary to assist newborns and people requiring support. Simulators for school fires, earthquakes, and flooding allow users to confirm evacuation routes inside schools, which are often designated as evacuation sites. Some simulators also allow users to experience routes blocked by fire or to evacuate together with other NPCs (Matsumura et al., 2023). However, because these simulators

are developed to reproduce individual disaster types, they do not reproduce route blockages caused by the earthquake that may precede and trigger tsunami-related flooding.
Simulators have also been developed in which NPCs that deliberately do not evacuate or NPCs that evacuate correctly are placed in the field (Kinateder & Warren, 2016; Matsumura et al., 2023; Mitsuhara et al., 2019). These studies mainly target fire scenes and earthquake disasters and have reported that participants may delay the transition to evacuation behavior or fail to evacuate by aligning their actions with those of NPCs, indicating conformity bias. During a disaster, conditions differ from ordinary times, and some evacuees are likely to be confused. If surrounding evacuees conform to such confused individuals, they may take incorrect evacuation actions.

In addition, psychological studies have examined collective behavior in crisis situations (Deutsch & Gerard, 1955; Drury et al., 2009; Mawson, 2005). From these perspectives, when people perceive danger, their decisions may be delayed as they conform to surrounding people, they may attempt to gather with close others such as family members or acquaintances, and a sense of solidarity may emerge from a shared perception of threat. In emergencies, collective behavior differs from that in ordinary situations. Both life-threatening behaviors and helping behaviors may therefore be rooted in human tendencies to maintain connections with others.

Taken together, prior studies show that simulators reproducing various disasters have been developed for both indoor and outdoor settings. However, two issues remain. The first is that existing simulators have not sufficiently reproduced the complex and irrational behaviors observed in actual disasters. During a disaster, evacuees in a state of panic may choose incorrect routes, and conformity bias toward such evacuees may lead to evacuation failure; however, this issue has not been clarified in the studies reviewed above. The second issue is that indoor flooding simulators have not reproduced route blockages. In tsunami-related flooding, an earthquake is likely to have occurred immediately beforehand, and a school building serving as an evacuation site may be blocked by collapsed objects caused by the earthquake. Existing flooding simulators do not reproduce such conditions. As a result, they do not account for situations in which evacuees assume that they can continue using the same staircase from the ground floor to reach the rooftop, or in which routes taught during ordinary times are no longer available.

To address the first issue, namely the inability to reproduce complex and irrational behavior observed in actual disasters, this study places multiple evacuees moving in incorrect directions, thereby reproducing irrational behavior and examining whether participants conform to it. To address the second issue, namely the absence of blockages in indoor flooding simulators, the indoor field is flooded while important routes for vertical evacuation, such as some staircases, are blocked. This design enables participants to learn multiple evacuation routes.

2.2 Overview of Methodology

Based on the issues discussed above, this study developed a simulator that floods a school, which is likely to serve as an evacuation site, and randomly places multiple NPCs inside the school. Some NPCs evacuate in the wrong direction, whereas others evacuate in the correct direction. Figure 1 shows an overview of the developed simulator. The proposed content consists of a flood-damage reproduction function and an evacuation-NPC reproduction function. Participants begin evacuation from the school gate, which is the Start Point shown in Figure 1. After exploring the school, evacuation is completed when they reach the Goal Point, the rooftop at the highest location. In this study, VR does not refer to immersive VR using a head-mounted display. Rather, it refers to a three-dimensional virtual space displayed on a laptop computer. Participants move through the virtual space using a keyboard and mouse in an ordinary

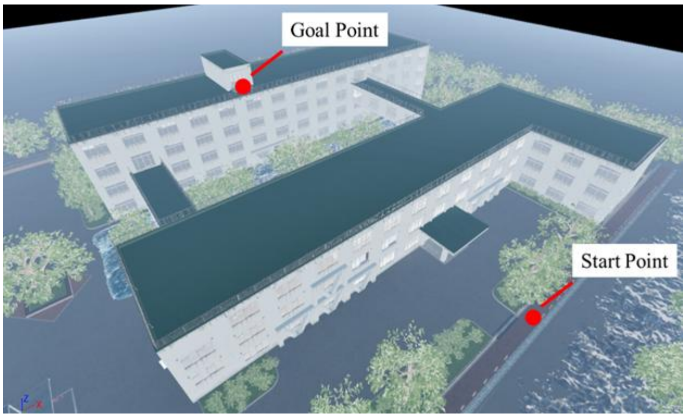

Figure 1. Overview of the developed simulator

PC operating environment and perform vertical evacuation while observing the flooding conditions, blocked areas, and NPC evacuation behavior.

2.3 Flood Damage Simulation

This function uses fluid simulation to construct a school field that is gradually flooded while placing obstacles on some staircases, thereby reproducing a more realistic flooding scenario. The 3D school model was created using the "Japanese School - Mega Pack with Full School Environment & 12 asset packs" (2025), available through Fab for Unreal Engine. When constructing the 3D model, branching routes were designed so that the lower portions of staircases could be seen when viewed closely. The field consists of two school buildings, one with three floors and one with four floors, and a rooftop. This structure was designed so that participants could find the correct route if they carefully explored it. For the water-flow simulation, Niagara Fluids (2025), shown in Figure 2, was used so that the school gradually became inundated. The flooding speed was set so that, if participants chose an incorrect route and reached a dead end blocked by collapsed objects, returning to the correct route would be difficult. To make participants experience the urgency of flooding, a time limit was displayed in the upper-left corner of the screen, and the trial ended in Game Over when the time limit expired. Even if the time limit had not expired, evacuation was considered extremely difficult once the participant was submerged up to the knees. Therefore, Game Over was triggered if the water level remained at or above the participant's knees for 10 s. In many schools, staircases continue from the ground floor to upper floors at the same location. During a disaster, however, such staircases may be blocked, or congestion on a blocked staircase may cause evacuees to become trapped. Therefore, as shown in Figure 3, this function reproduces blockages caused by collapsed objects on staircases and congestion caused by NPCs. This function is expected to help address the limitation that indoor flooding simulators have not reproduced route blockages.

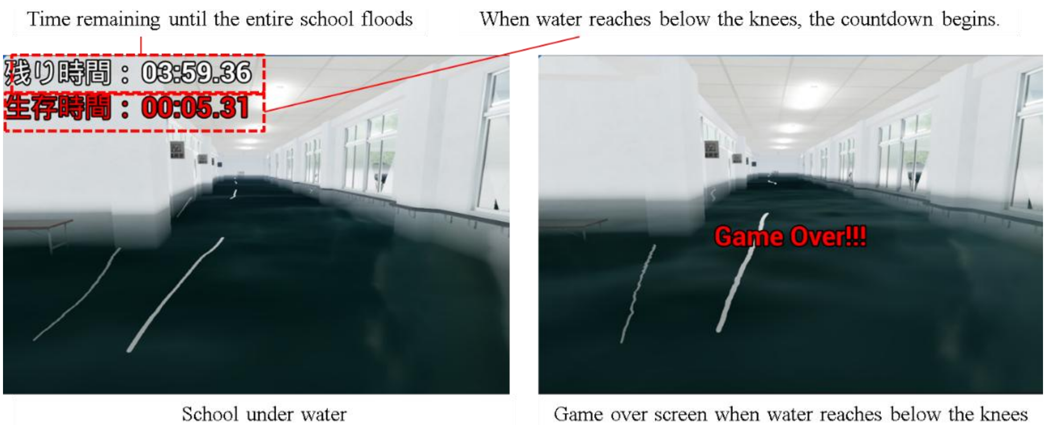

Figure 2. Recreation of Flood Damage

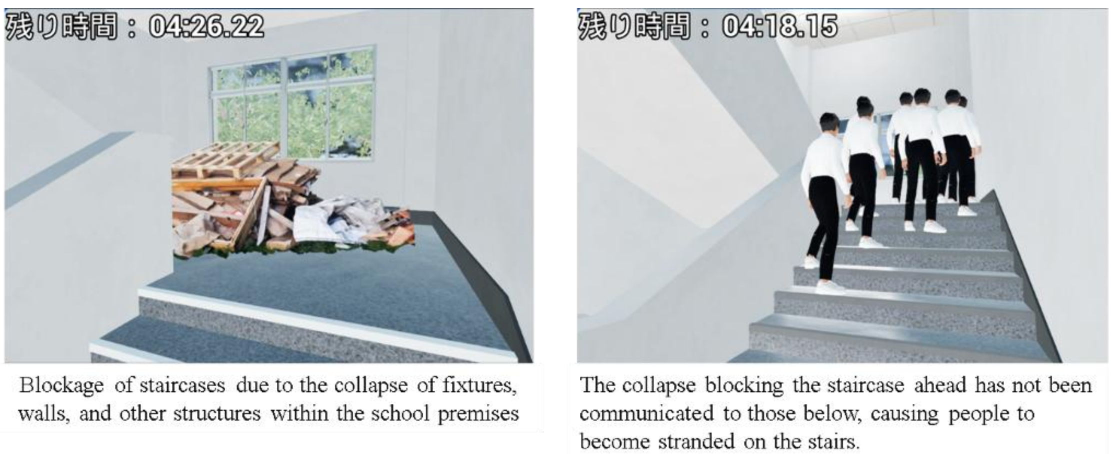

Figure 3. Blockage of the staircase

2.4 Evacuation NPC Reproduction

This function populates the corridors of each building with other evacuees to stimulate participants' conformity bias. From the perspective of normative and informational influence (Deutsch & Gerard, 1955), the behavior of others may affect decision making when judgment is difficult. Studies of collective behavior during disasters have also indicated that people's actions in crisis situations are not simply panic responses but are influenced by relationships with others and interactions within groups (Drury et al., 2009; Mawson, 2005). Accordingly, this study established two NPC conditions: Random NPCs, which evacuate randomly, and Conformity NPCs, which include NPCs that move toward incorrect routes. Examples of the reproduced NPCs are shown in Figure 4. Random NPCs were prepared because they also move in directions that are clearly not evacuation routes and were therefore expected to be less likely to induce conformity bias. By contrast, the Conformity NPC condition included both NPCs that select correct evacuation routes and NPCs that select incorrect routes and then stop after reaching a dead end. This function makes it possible to examine whether the irrational behavior of other evacuees affects participants' evacuation behavior and psychological state. It is therefore expected to help address the limitation that existing simulators have not reproduced complex and irrational behavior observed in actual disasters.

2.5 Experimental System and Protocol

Figure 5 shows the experimental procedure used in this study. The experiment did not use a VR headset; instead, it used a three-dimensional virtual space displayed on a laptop computer. The experimental system was built with Unreal Engine 5. The display device was the laptop's built-in 14.5-inch OLED display with a resolution of 2880 x 1800 pixels. The simulation was run on a laptop equipped with an Intel Core i9-13900H CPU, an NVIDIA GeForce RTX 4070 GPU, and 32 GB of main memory.

Participants completed the experiment in an ordinary PC operating environment. Keyboard input was used for movement, and mouse operation was used to control the viewing direction. Participants could move forward, backward, left, and right using the WASD keys or arrow keys, and movement speed was fixed at a constant value. Viewpoint

NPCs evacuating randomly　　NPC evacuating to the wrong route

Figure 4. NPCs simulating evacuation behavior

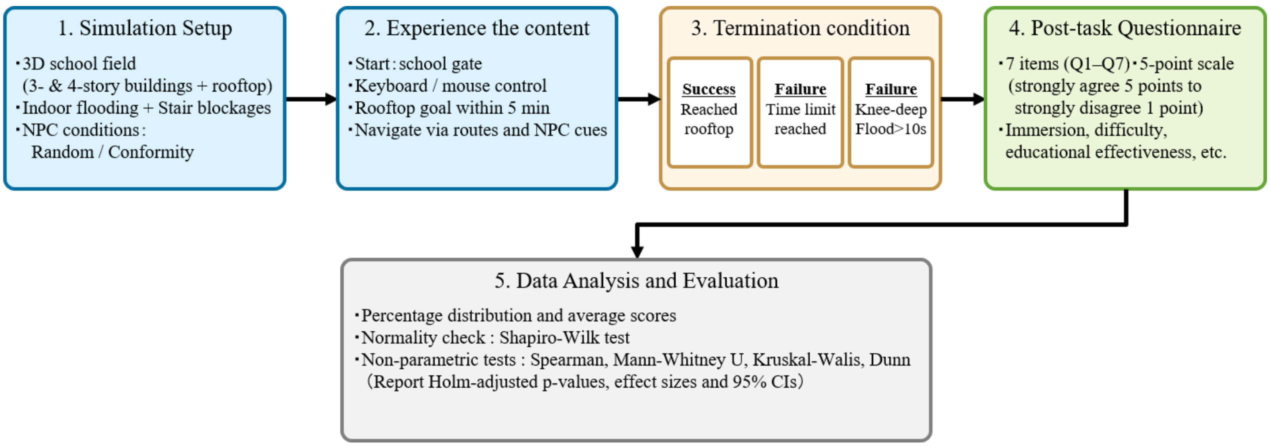

Figure 5. Experimental procedure

sensitivity was also standardized across all participants. No in-game tutorial was provided. Before the trial began, the experimenter explained how to move and control the viewpoint, confirmed that the participant had acquired the basic operations, and then started the main trial.

In the virtual space, NPCs capable of movement similar to that of the participant were placed as a crowd moving around the participant. In both conditions, 40 NPCs were placed. In the Conformity NPC condition, the 40 NPCs were divided at a 1:1 ratio between those selecting the correct evacuation route and those selecting an incorrect route. NPC movement speed was set to be equivalent to that of the participant. To enable participants to observe NPC behavior immediately after the experiment began, NPCs were initially placed in corridor locations. They were arranged at regular intervals to prevent excessive overlap among NPCs.

As a rule for route selection, NPCs were designed to move along routes set in advance. If a participant or another NPC was present on the route, the NPC temporarily stopped and resumed movement once passage became possible. NPCs selecting the correct route were assigned routes that led to the destination, whereas NPCs selecting an incorrect route were assigned routes leading in different directions. In contrast, participants could move freely through the space and made evacuation decisions while observing the behavior of NPCs.

Each trial had a time limit of 5 min, and participants were instructed to reach the rooftop, the destination point, within that time. If time expired during the trial, or if 10 s elapsed after the participant was submerged up to the knees, the trial was treated as an evacuation failure. If a system problem occurred during the experiment, such as application stoppage or loss of operability, or if unexpected congestion among NPCs made it completely impossible for the participant to move, the trial was immediately interrupted. When a system problem occurred, the experimental environment was restarted and the participant repeated the trial under the same condition. If unexpected movement failure occurred because of NPC congestion, the participant also repeated the trial under the same condition; however, becoming stuck at intentionally blocked routes was not treated as grounds for a repeat trial. Participants were also informed that they could stop the experiment at any time if they found it difficult to continue for psychological or physical reasons.

3. Results

3.1 Experimental Conditions

The primary purpose of this experiment was to clarify how the developed simulator was perceived as educational material for experiencing vertical evacuation during indoor flooding. After experiencing the simulator, participants completed a questionnaire survey to examine subjective evaluations and differences in evaluation tendencies associated with NPC evacuation-behavior conditions. The questionnaire items are shown in Table 1. Participants responded to seven items, Q1 to Q7, on a five-point scale for each item, ranging from strong agreement (5 points) to strong disagreement (1 point). The participants consisted of 60 people in total: 12 high school students living in Osaka Prefecture, 27 university students, and 21 visitors to the National Conference for Promoting Disaster Risk Reduction 2025, who were presumed to have relatively high disaster preparedness awareness. The high school students were students interested in the information field who attended an open campus event at an information-oriented university. The university students were students studying at an information-oriented university. Each participant experienced only one of the two simulators: one containing only Random NPCs or one containing only Conformity NPCs. The Random NPC condition included 12 high school students (Group 1) and 18 university students (Group 2), whereas the Conformity NPC condition included 21 visitors to the National Conference for Promoting Disaster Risk Reduction

Table 1. Questionnaire Items

Question	Detail
Q1	Do you usually play games?
Q2	Do you take disaster preparedness measures when you are in your usual living areas such as at home or school?
Q3	Do you take disaster prevention measures when you are out and about, or when you are "outside your usual living area"?
Q4	The simulation felt realistic and immersive
Q5	Did you find the simulation difficult, such as getting lost?
Q6	Did you feel any sense of crisis or impatience during the simulation?
Q7	Did this experience make you more aware of disaster prevention?

2025 (Group 3) and nine university students (Group 4). Information on sex or gender was not collected. This decision was made to protect participant privacy, to respect diverse gender identities, and to collect only the minimum personal attributes necessary for the research purpose.

This study did not randomly assign participants from the same population to the Random NPC and Conformity NPC conditions. Therefore, the NPC condition and participant background were not fully independent, and differences between NPC conditions may reflect not only differences in NPC evacuation behavior but also differences in participant attributes. For this reason, differences between NPC conditions are interpreted as tendencies in subjective evaluations rather than as pure NPC effects.

3.2 Experimental Results

3.2.1 Statistical Analysis

For all statistical tests, the significance level was set at 5%. For analyses involving multiple comparisons, p-values were adjusted using the Holm method. The 95% confidence intervals for effect sizes were standard, unadjusted intervals. Effect sizes were reported as Spearman’s rho for Spearman’s rank correlation coefficient, rank-biserial correlation for the Mann-Whitney U test, and rank eta-squared for the Kruskal-Wallis test. In the Shapiro-Wilk test, the null hypothesis was that the response distribution for each item followed a normal distribution, and the alternative hypothesis was that the response distribution did not follow a normal distribution. In Spearman’s rank correlation analysis, the null hypothesis was that there was no monotonic association between the two questionnaire items, and the alternative hypothesis was that there was a monotonic association. In the Mann-Whitney U test, the null hypothesis was that the response distributions for the Random NPC and Conformity NPC conditions did not differ, and the alternative hypothesis was that the response distributions differed between the two conditions. In the Kruskal-Wallis test, the null hypothesis was that the response distributions did not differ among the four participant groups, and the alternative hypothesis was that at least one group differed from the others. In Dunn’s test, the null hypothesis was that the response distributions did not differ between the two groups being compared, and the alternative hypothesis was that they differed.

Because the questionnaire items Q1 to Q7 were measured on a five-point ordinal scale, the normality of the distribution of each item was first examined using the Shapiro-Wilk test. The results are shown in Table 2. As shown in Table 2, normality was not confirmed for any item. Therefore, this study used nonparametric tests for the analyses.

3.2.2 Analysis for all participants

The overall response results for all participants are shown in Table 3. As shown in Table 3, the average scores for Q4 and Q7 were 4 or higher, whereas only Q3 had an average score close to 2. These results indicate that, although many participants did not usually take disaster preparedness measures when going outside their usual living areas, the simulator enabled them to experience realistic vertical evacuation while avoiding blockages in the presence of other evacuees and encouraged them to reflect on disaster preparedness. Thus, the simulator addresses the limitation of prior indoor flooding simulators that did not reproduce blockages.

Table 2. Results of the Shapiro-Wilk Test

	Q1	Q2	Q3	Q4	Q5	Q6	Q7
W	0.732	0.865	0.78	0.726	0.814	0.750	0.802
p	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**

Bold : p<0.05

Table 3. Overall Response Results for All Participants

Score	Q1	Q2	Q3	Q4	Q5	Q6	Q7
5	51.7%	15.0%	5.0%	48.3%	31.7%	46.7%	40.0%
4	21.7%	33.3%	15.0%	41.7%	40.0%	30.0%	41.7%
3	6.7%	6.7%	3.3%	3.3%	3.3%	5.0%	13.3%
2	6.7%	26.7%	35.0%	6.7%	18.3%	8.3%	3.3%
1	13.3%	18.3%	41.7%	0.0%	6.7%	10.0%	1.7%
Average Score	3.91	3.00	2.07	4.31	3.72	3.95	4.15

Next, associations among Q1 to Q7 for all participants were analyzed using Spearman's rank correlation coefficient. The results are shown in Table 4. Significant positive correlations were observed for Q2-Q3 and Q5-Q6. For Q2-Q3, participants who took disaster preparedness measures in their usual living areas were also more likely to take measures when going outside those areas. For Q5-Q6, participants who perceived the simulation as difficult tended to report a stronger sense of crisis or impatience.

Table 4. Results of Spearman's Rank Correlation Coefficient

Question item pairs	Spearman's ρ	p holm	Lower 95% CI	Upper 95% CI	Fisher's z
Q1-Q2	-0.144	1	-0.384	0.114	-0.145
Q1-Q3	-0.374	0.057	-0.573	-0.132	-0.393
Q1-Q4	0.118	1	-0.140	0.361	0.119
Q1-Q5	-0.215	1	-0.445	0.041	-0.218
Q1-Q6	-0.312	0.27	-0.524	-0.063	-0.322
Q1-Q7	-0.092	1	-0.338	0.165	-0.093
Q2-Q3	0.586	**<.001**	0.390	0.731	0.671
Q2-Q4	0.067	1	-0.190	0.316	0.067
Q2-Q5	0.031	1	-0.225	0.283	0.031
Q2-Q6	0.112	1	-0.146	0.356	0.113
Q2-Q7	0.277	0.512	0.025	0.496	0.285
Q3-Q4	0.057	1	-0.200	0.307	0.057
Q3-Q5	0.078	1	-0.179	0.326	0.078
Q3-Q6	0.208	1	-0.049	0.439	0.211
Q3-Q7	0.137	1	-0.121	0.378	0.138
Q4-Q5	0.114	1	-0.144	0.358	0.114
Q4-Q6	0.242	0.93	-0.012	0.468	0.247
Q4-Q7	0.150	1	-0.108	0.39	0.152
Q5-Q6	0.541	**<.001**	0.333	0.699	0.605
Q5-Q7	0.282	0.493	0.030	0.500	0.290
Q6-Q7	0.168	1	-0.090	0.404	0.169

Bold : p<0.05

Table 5. Survey Results by NPC Type

NPC Type	Score	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Random (N=30)	5	60.0%	6.7%	0.0%	53.3%	16.7%	33.3%	36.7%
	4	26.7%	30.0%	13.3%	33.3%	43.3%	40.0%	36.7%
	3	6.7%	3.3%	0.0%	3.3%	0.0%	0.0%	16.7%
	2	6.7%	30.0%	26.7%	10.0%	26.7%	13.3%	6.7%
	1	0.0%	30.0%	60.0%	0.0%	13.3%	13.3%	3.3%
	Average Score	4.40	2.53	1.67	4.30	3.23	3.66	3.97
Conformity (N=30)	5	43.3%	23.3%	10.0%	43.3%	46.7%	60.0%	43.3%
	4	16.7%	36.7%	16.7%	50.0%	36.7%	20.0%	46.7%
	3	6.7%	10.0%	6.7%	3.3%	6.7%	10.0%	10.0%
	2	6.7%	23.3%	43.3%	3.3%	10.0%	3.3%	0.0%
	1	26.7%	6.7%	23.3%	0.0%	0.0%	6.7%	0.0%
	Average Score	3.44	3.47	2.47	4.33	4.20	4.23	4.33
Mann-Whitney U test	U	579	279.5	269	474	263.5	332	376
	p holm	0.156	**0.045**	**0.028**	0.700	**0.028**	0.189	0.484
	r_{rb}	-0.287	0.379	0.402	-0.053	0.414	0.262	0.164
	Lower 95% CI	-0.5300	0.1040	0.1310	-0.3350	0.1450	-0.0260	-0.1280
	Upper 95% CI	-0.0003	0.6000	0.6170	0.2370	0.6260	0.5100	0.4310

Bold : p<0.05

3.2.3 Analysis by NPC type

To examine whether participant responses differed significantly according to NPC type, the results were analyzed using the Mann-Whitney U test. Table 5 shows the results. Looking first at the average scores, the Random NPC condition had scores of 4 or higher for Q1 and Q4, whereas the Conformity NPC condition had scores of 4 or higher for Q4, Q5, Q6, and Q7. Thus, the Conformity NPC condition tended to have high average scores for more items. Because Q4 was 4 or higher regardless of NPC type, differences in NPC evacuation behavior appear not to have reduced the perceived realism of the simulator. The Mann-Whitney U test showed significant differences between conditions for Q2, Q3, and Q5. However, because participant attributes differed between NPC types, some of these significant differences may have resulted from differences in participant background. Therefore, group-based analysis was also necessary.

3.2.4 Analysis of Participants by Group

The group-specific response results are shown in Table 6. Because the questionnaire survey included four participant groups, the Kruskal-Wallis test was used to examine whether significant differences existed among three or more groups. The results are shown in Table 7. Significant differences were found for Q1, Q2, and Q3. Because the Kruskal-Wallis test alone does not identify which groups differ, Dunn's test was then conducted for each item in which a significant difference was found. The results are shown in Table 8. In Q1 and Q3, significant differences were observed for the Group 1-Group 3 pair. In Q2, significant differences were observed for the Group 2-Group 3 and Group 3-Group 4 pairs.

Table 6. Group-Specific Response Results

NPC Type	Group	Score	Q1(%)	Q2(%)	Q3(%)	Q4(%)	Q5(%)	Q6(%)	Q7(%)
Random (N=30)	Group1 (N=12)	5	91.7	16.7	0.0	83.3	33.3	41.7	41.7
		4	8.3	33.3	8.3	8.3	25.0	33.3	41.7
		3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		2	0.0	16.7	8.3	8.3	16.7	0.0	8.3
		1	0.0	33.3	83.3	0.0	25.0	25.0	8.3
	Group2 (N=18)	5	38.9	0.0	0.0	33.3	5.6	27.8	33.3
		4	38.9	27.8	16.7	50.0	55.6	44.4	33.3
		3	11.1	5.6	0.0	5.6	0.0	0.0	27.8
		2	11.1	38.9	38.9	11.1	33.3	22.2	5.6
		1	0.0	27.8	44.4	0.0	5.6	5.6	0.0
Conformity (N=30)	Group3 (N=21)	5	33.3	33.3	14.3	42.9	42.9	61.9	61.9
		4	14.3	42.9	23.8	52.4	42.9	19.0	33.3
		3	9.5	9.5	9.5	4.8	9.5	14.3	4.8
		2	4.8	9.5	38.1	0.0	4.8	4.8	0.0
		1	38.1	4.8	14.3	0.0	0.0	0.0	0.0
	Group4 (N=9)	5	66.7	0.0	0.0	44.4	55.6	55.6	0.0
		4	22.2	22.2	0.0	44.4	22.2	22.2	77.8
		3	0.0	11.1	0.0	0.0	0.0	0.0	22.2
		2	11.1	55.6	55.6	11.1	22.2	0.0	0.0
		1	0.0	11.1	44.4	0.0	0.0	22.2	0.0

Table 7. Results of the Kruskal-Wallis Test

Question	H	p holm	Rank η^2	Lower 95% CI	Upper 95% CI
Q1	14.73	**0.012**	0.209	0.082	0.427
Q2	14.41	**0.012**	0.204	0.051	0.479
Q3	16.53	**0.006**	0.242	0.062	0.508
Q4	6.438	0.184	0.061	0	0.306
Q5	8.888	0.092	0.105	0.023	0.37
Q6	4.000	0.261	0.018	0	0.233
Q7	9.662	0.086	0.119	0.026	0.344

Bold : p<0.05

Table 8. Results of Dunn's Test

Question	Group pair	z	r_{rb}	p holm
Q1	Group1-Group2	2.325	-0.546	0.100
	Group1-Group3	3.655	-0.627	**0.002**
	Group1-Group4	0.971	-0.259	0.566
	Group2-Group3	1.420	-0.307	0.467
	Group2-Group4	-1.074	0.272	0.566
	Group3-Group4	2.245	-0.466	0.100
Q2	Group1-Group2	0.954	-0.157	1
	Group1-Group3	-2.126	0.373	0.134
	Group1-Group4	0.593	-0.093	1
	Group2-Group3	-3.502	0.640	**0.003**
	Group2-Group4	-0.230	0.093	1
	Group3-Group4	-2.588	0.646	**0.048**
Q3	Group1-Group2	-1.696	0.370	0.270
	Group1-Group3	-3.858	0.702	**<.001**
	Group1-Group4	-1.041	0.343	0.596
	Group2-Group3	-2.379	0.431	0.087
	Group2-Group4	0.423	-0.093	0.672
	Group3-Group4	-2.352	0.566	0.087

Bold : p<0.05

4. Discussion

4.1 Discussion of Overall Results

This section discusses the overall tendencies among participants. Table 3 shows that, for Q7, the combined proportion of strong and weak agreement was 81.7%, indicating that many participants subjectively reported an increase in disaster preparedness awareness after the experience. However, because Q7 was a single post-experience self-report item, this result alone cannot directly demonstrate an improvement in disaster preparedness awareness. Rather, the result suggests that the developed simulator may provide an opportunity to raise awareness of vertical evacuation.

4.2 Analysis of Results by NPC Type

This section discusses tendencies by NPC type. First, the Q5 results in Table 5 show that participants in the Conformity NPC condition rated the simulation as more difficult than those in the Random NPC condition, and the difference was statistically significant after Holm correction. In the Conformity NPC condition, participants encountered not only NPCs selecting correct routes but also NPCs moving toward incorrect routes. Participants therefore needed to judge their evacuation route while observing the behavior of other evacuees. This may have caused the Conformity NPC condition to be perceived as a more difficult evacuation experience than the Random NPC condition.

These results suggest that the developed simulator may have educational value by addressing the limitation that previous studies did not reproduce complex and irrational behaviors observed in actual disasters. However, no statistically significant differences were found for Q6 or Q7. This suggests that differences in NPC movement alone may not be sufficient to produce a clear conformity-bias effect or a large difference in educational outcomes. Significant differences were also found for Q2 and Q3, but these items concern participants' own disaster preparedness behavior. Because the NPC condition and participant background were not fully independent in this study, those differences are likely to reflect participant characteristics rather than NPC effects.

From the perspective of informational influence described by Deutsch and Gerard (1955), the behavior of others may serve as a cue for decision making in uncertain situations. In the Conformity NPC condition, placing NPCs that moved toward incorrect routes may have increased participants' uncertainty in judging evacuation routes, leading to higher ratings of subjective difficulty. However, because this study did not collect behavioral indices such as the NPC-following rate or movement trajectories, these results are not direct evidence of conformity behavior itself. They should instead be interpreted as indicating the possibility that others' behavior affected subjective difficulty.

4.3 Discussion of Results by Participant Group

This section discusses the group-based analysis that could not be fully addressed in the NPC-type analysis. The results for Q2 and Q3 in Table 8 show that significant differences were found mainly between Group 3 and the other groups. The descriptive values in Table 6 also indicate that Group 3, consisting of participants from the National Conference for Promoting Disaster Risk Reduction 2025, had higher disaster preparedness awareness than the other groups. In this experiment, a group with higher disaster preparedness awareness experienced only the Conformity NPC condition. Therefore, it is difficult to evaluate differences in educational effects as being due purely to differences in NPC type. For Q7, the Kruskal-Wallis test did not reach statistical significance after Holm correction. Therefore, this study does not claim differences in educational effects based on group differences. Although the descriptive statistics show high ratings in some groups, participants' prior disaster preparedness awareness and background may have influenced their evaluations. Future studies should use pre- and post-measurements, random assignment within the same population, within-subject comparisons, or analyses that control for participant attributes to verify more rigorously how this simulator affects disaster preparedness awareness.

4.4 Limitations and Future Work

This study has several limitations. First, the evaluation was based on subjective ratings of Q1 to Q7 on a five-point scale. The study did not collect or analyze objective behavioral logs such as movement trajectories, the number of incorrect route choices, the NPC-following rate, or stopping time. Therefore, the present results alone cannot demonstrate directly that participants conformed to others and took incorrect actions, nor can they establish the direct occurrence of conformity bias. Future work should extend the content so that such behavioral logs can be collected, enabling quantitative analysis of how incorrect NPC behavior affects participants' route choices and stopping behavior.

Second, participant backgrounds were not fully independent between the Random NPC and Conformity NPC conditions. Differences observed between conditions may therefore reflect not only differences in NPC evacuation behavior but also participant attributes. Future studies should use random assignment within the same population, within-subject comparisons, or analyses that control for participant attributes to examine the effects of NPC behavior more rigorously.

Third, this study was an empirical evaluation conducted through events and educational opportunities, and the sample size was not determined by an a priori power analysis. In particular, the group-based analysis involved small and unbalanced sample sizes, limiting statistical power. Therefore, nonsignificant results do not necessarily indicate the absence of an effect, and significant results should also be interpreted with caution. The difference in subjective difficulty between NPC conditions was an analysis anticipated in advance, whereas correlations among questionnaire items and group-based comparisons should be regarded as supplementary and exploratory analyses.

5. Conclusions

This study developed a simulator for vertical evacuation from flooded indoor areas to a rooftop. The simulator targets schools, which are often designated as evacuation sites, and incorporates NPCs that may induce conformity bias as well as route blockages caused by collapse. A questionnaire survey of 60 participants was conducted to evaluate the simulator's usefulness. Many participants reported that the simulator felt realistic and immersive. In addition, the Conformity NPC condition tended to be rated as subjectively more difficult than the Random NPC condition. The analysis of associations among questionnaire items further showed that participants who perceived the simulation as difficult tended to report a stronger sense of crisis or impatience. These results suggest that the simulator may be useful as a disaster preparedness education tool that enables participants to experientially consider the difficulty and urgency of vertical evacuation during indoor flooding.

At the same time, this study did not collect or analyze objective behavioral logs such as movement trajectories, the number of incorrect route choices, the NPC-following rate, or stopping time. It therefore cannot directly demonstrate the occurrence of conformity bias or incorrect evacuation behavior caused by conformity to NPCs. In addition, because the NPC condition and participant background were not fully independent, differences between conditions cannot be attributed solely to differences in NPC behavior. Future studies should collect behavioral logs and use random assignment within the same population, within-subject comparisons, or analyses that control for participant attributes to examine more rigorously how NPC evacuation behavior affects participants' evacuation decisions and how the simulator contributes to disaster preparedness education. Through these improvements and reevaluation, we aim to more rigorously examine the simulator's potential educational effectiveness.

Author Contributions

Conceptualization, M.N., T.K. and M. K.; methodology, M.N., T.K. and M.K.; software, T.K.; validation, M.N., T. K. and M.K.; formal analysis, T.K.; investigation, T.K. and M. K.; resources, Y.T., Y.K., and Y.A.; data curation, M.N.; writing—original draft preparation, T.K.; writing—review and editing, M.N., Y.T., Y.U., Y.T., Y.K., Y.A. and M.K.; visualization, T.K. and M.K.; supervision, M.N., Y.T. and Y.U.; project administration, M.N.

Funding

This research was conducted with funding from a commissioned study by the National Institute of Information and Communications Technology (Grant No. 23603) and JSPS KAKENHI Grant Number JP26K01183. We express our gratitude here.

Institutional Review Board Statement

This research was reviewed and approved by the Bioethics Committee of Osaka Electro-Communication University and the Faculty Ethics Committee of Reitaku University.

Informed Consent Statement

Participants were informed that the survey results might be used for academic presentations and publications, and consent was obtained prior to participation.

Conflicts of Interest

The authors declare no conflict of interest.

References:

Arakawa, T., Obayashi, F., Kobayashi, K., Itamiya, T., Uno, S., Yamabe, S., Suzuki, T. & Tajima, J. (2025). Simulation-based tsunami evacuation training framework aimed at avoiding the negative consequences of using cars, *Geosciences*, 15(9), 364. https://doi.org/10.3390/geosciences15090364

Deutsch, M. & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. The Journal of Abnormal and Social Psychology, 51(3), 629-636. https://doi.org/10.1037/h0046408

Drury, J., Cocking, C. & Reicher, S. (2009). The nature of collective resilience: Survivor reactions to the 2005 London bombings. International Journal of Mass Emergencies and Disasters, 27(1), 66-95. https://doi.org/10.1177/028072700902700104

Farra, S., Hodgson, E., Miller, T.E., Timm, N., Brady, W., Gneuhs, M., Ying, J., Hausfeld, J., Cosgrove, E., Simon, A. & Bottomley, M. (2018). Effects of virtual reality simulation on worker emergency evacuation of neonates, *Disaster Medicine and Public Health Preparedness, 13*(2), pp.301-308. https://doi.org/10.1017/dmp.2018.58

Iinuma, K., Kobayashi, K., Shinozaki, I., Shimazaki, Y., Chiba, T., Yamabe, Y., Moriya, T., Kodama, S. & Takahashi, T. (2024). VR System for Evacuation Experience from Tsunami Flooding, *ITE Tech. Rep., 48*(8), pp.366-369.

Japan Meteorological Agency. (2025). Retrieved October 25, 2025, from https://www.jma.go.jp/jma/kishou/know/jishin/nteq/assumption.html

Japanese School - Mega Pack with Full School Environment & 12 asset packs. (2025). Retrieved October 25, 2025, from https://www.fab.com/pt-br/listings/93bd3af2-1edd-4a99-a4a7-dc4b7bda3c87

Kaggwa, M. M., Chaimowitz, A. G., Agboinghale, P., Marginean, D. & Olagunju, T. A. (2025). Virtual reality training programs in disaster preparedness: a systematic review, *Discover Education, 4*(1), pp.325. https://doi.org/10.1007/s44217-025-00771-5

Kinateder, M. & Warren, W. (2016). Social influence on evacuation behavior in real and virtual environments, *Front. Robot. AI, 3*. https://doi.org/10.3389/frobt.2016.00043

Lorusso, P., Iuliis, M., Marasco, S., Domaneschi, M., Cimellaro, G. & Villa, V. (2022). Fire emergency evacuation from a school building using an evolutionary virtual reality platform, *Buildings, 12*(2), 223. https://doi.org/10.3390/buildings12020223

Marui, T., Wada, M., Fumoto, Y., Fujimori, Y., Shiba, D., Hatori, T., Matsumura, N. & Moriwaki, R. (2024). Development of VR for tsunami evacuation and evaluation of its effectiveness in disaster prevention education targeting local students, *Journal of disaster information studies, 22*, pp.69-78.

Matsumura, N., Sugihara, K. & Yonezawa, T. (2023). Effect of agents' number and behaviors in VR disaster evacuation drill, *IEICE Tech. Rep.*, 123(60), pp.13-18.

Mawson, A. R. (2005). Understanding mass panic and other collective responses to threat and disaster. Psychiatry, 68(2), 95-113. https://doi.org/10.1521/psyc.2005.68.2.95

Ministry of Land, Infrastructure, Transport and Tourism. (2024). Retrieved October 25, 2025, from https://www.mlit.go.jp/river/bousai/240101_noto/pdf/tsunamishinsui_higai_240423.pdf

Mitsuhara, H. (2024). Metaverse-based evacuation training: design, implementation, and experiment focusing on earthquake evacuation, *Multimodal Technol. Interact., 8*(12), 112. https://doi.org/10.3390/mti8120112

Mitsuhara, H., Tanimura, C., Nemoto, J. & Shishibori, M. (2019). Failure-enhanced evacuation training using a VR-based disaster simulator: A comparative experiment with simulated evacuees, *Procedia Computer Science, 159*, pp.1670-1679. https://doi.org/10.1016/j.procs.2019.09.337

Nakai, H., Itatani, T., Horiike, R., Kyota, K. & Tsukasaki, K. (2018). Tsunami evacuation simulation using geographic information systems for homecare recipients depending on electric devices, *PLoS ONE, 13*(6), e0199252. https://doi.org/10.1371/journal.pone.0199252

National Federation of Workers and Consumers Kyosai Cooperatives. (2024). Retrieved October 25, 2025, from https://www.zenrosai.coop/bousai-gensai/survey2024.html

Niagara Fluids, (2025). Retrieved October 25, 2025, from https://dev.epicgames.com/documentation/ja-jp/unreal-engine/niagara-fluids-in-unreal-engine

Shimbo, T., Terayama, K., Koshino, M., Okino, K., Araki, K. & Yoshita, R. (2022). Development of “disaster prevention sugoroku,” an application for disaster prevention education using VR content, and its educational effects, *Journal of Japan Society of Civil Engineers, Ser. H (Engineering Education and Practice)*, *78*(1) pp.1-9. https://doi.org/10.2208/jscejeep.78.1_1