2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) | 978-1-6654-5365-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISMAR-ADJUNCT57072.2022.00097

Synthesizing Shared Space Virtual Reality Fire Evacuation Training Drills

Huimin Liu[1*] Minsoo Choi[1†] Liuchuan Yu[2‡] Alexandros Koilias[3§] Lap-Fai Yu[2¶] Christos Mousas[1‖]

[1]Department of Computer Graphics Technology, Purdue University, West Lafayette, Indiana 47907, U.S.A.
[2]Department of Computer Science, George Mason University, Fairfax, Virginia 22030, U.S.A.
[3]Department of Cultural Technology and Communication, University of the Aegean, Mytilene 81100, Greece

Abstract

We synthesized virtual reality fire evacuation training drills in a shared virtual space to explore people's collaboration behavior. We formulate the authoring process of the fire evacuation training drill in a total cost function, which we later solve with a Markov Chain Monte Carlo (MCMC) optimization-based method. The users' assigned task in the synthesized training drill is to help virtual agents evacuate the building as quickly as possible using predefined interaction mechanisms. The users can join the training drill from different physical locations and collaborate and communicate in a shared virtual space to finish the task. We conducted a user study to collect both in-game measurements and subjective ratings to evaluate whether the synthesized training drills would affect how the participants collaborated.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Software and its engineering—Contextual software domains—Virtual worlds software—Virtual worlds training simulations

1 Introduction

Collaboration is usually characterized by shared goals, group activities, communication, and exchanging information [17]. Roschelle and Teasley [26] defined collaboration on a joint problem space as the "mutual engagement of people in a coordinated effort to solve a problem together." Various researchers [5, 8] regard collaboration as an essential component of effective training and learning in comparison to individual tasks. In the age of fast-paced development of globalization, which has a higher requirement for productivity, especially during the COVID-19 pandemic when people have been impeded from meeting in person, the importance of remote collaboration systems has been emphasized, as they contribute to remote team task success, reduce travel expenses, ensure safety, reduce carbon emissions, increase efficiency, and save time and energy.

However, the concept of collaboration is abstract and difficult to grasp [11], making it challenging to utilize in practical applications. When implementing collaborative training scenarios in virtual environments, designers usually manually build the contents according to their subjective experiences and intuition in order to trigger the intended behavior in participants. This process is tedious and time-consuming since it lacks a solid theory that supports the effectiveness of the designed content. To better support collaboration on common tasks among the involved group members, it is necessary to obtain a more precise understanding of collaboration and how to conduct immersive collaboration remotely in a shared virtual space using modern virtual reality (VR) technologies.

*e-mail: liu2833@purdue.edu
†e-mail: choi714@purdue.edu
‡e-mail: lyu20@gmu.edu
§e-mail: ctd17008@aegean.gr
¶e-mail: craigyu@gmu.edu
‖e-mail: cmousas@purdue.edu

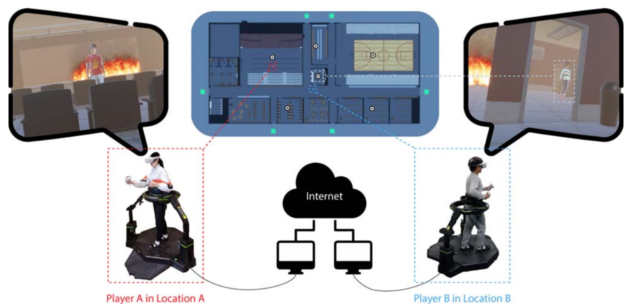


Figure 1: Two players in different locations, wearing a VR headset on the VR treadmill. Their task is to guide the agents out of the building where a simulated fire emergency occurs. The two players are in the same virtual space even though their physical locations are different. They can communicate, use voice commands to guide the agents outside the building, and use a fire extinguisher to eliminate the fire in the building. We illustrate users' and agents' positions and the top view of the building in the minimap.

The project presented in this paper focused on synthesizing VR fire evacuation training drills in a shared virtual space to explore the participants' collaboration behavior. Inspired by procedural content generation approaches, we proposed an optimization-based method that automatically generates fire evacuation training drills with varying levels of difficulty. The users' assigned task is to help virtual agents evacuate the building as quickly as possible using predefined interaction mechanisms (voice commands, trigger fire extinguisher, physical locomotion, etc.). The participants can join the training drill from different locations and collaborate and communicate in a shared virtual space to accomplish the task (see Fig. 1). We evaluated the proposed VR training drill authoring method by conducting a user study among three training drills with different difficulty levels: low difficulty (LD), medium difficulty (MD), and high difficulty (HD). We collected both in-game measurements and subjective ratings to explore how the participants collaborate in such a VR setup.

2 Related Work

Virtual reality (VR) and augmented reality (AR) are as effective of a training mechanism as the commonly accepted methods [15]. VR can enhance the learning and training. Some work focused on training for sports [20] and education [10]. Also, some research was conducted for medical and rehabilitation purposes [27], and for evacuation training and research purposes [19]. As for AR training, research shows that AR, applied in education and training, has positive potential for the future of education [18]. Moreover, AR shows great potentials and can be applied in many other fields, such as, medical education [3], corporate training [22], healthcare simulation [30], maintenance skills [32], and vocational training [6].

2771-1110/22/$31.00 ©2022 IEEE
DOI 10.1109/ISMAR-Adjunct57072.2022.00097

For more details about VR training, please refer to Xie et al. [34].

With network, VR and AR can be applied in remote training and collaboration scenarios. Greenwald et al. [13] explored the immense potential for collaborative VR applications for learning. Some researchers proposed frameworks to support collaboration in virtual environments. For example, MedicalVR [21] is a virtual reality framework and assistive tool for medical environment. It outlines real-time collaboration and human-centered design aspects in modern tele-medicine. Kurillo et al. [16] presented a framework for immersive virtual environment intended for remote collaboration and training of physical activities. For example, Tea et al. [29] developed a multi-user immersive virtual reality application for real-time remote collaboration to enhance design review process. Snow Dome [24], which is a mixed reality remote collaboration application, was developed to support multi-scale interaction for a virtual reality user. Elvezio et al. [9] demonstrated an approach to support remote collaboration in AR and VR by virtual replicas, which allows the remote user to create and manipulate virtual replicas of physical objects in the local environment. Besides from framework, system, and application, some research focused on adaptive avatar, Mini-Me [25], and toolkit, ColabAR [31], to promote remote collaboration.

In this paper, we propose an optimization-based method to automatically synthesize shared space VR fire evacuation training drills with different difficulty levels. We also demonstrated how to employ the synthesized training drills on a networked VR platform with treadmills to enable remote, collaborative training.

3 Preliminary Remarks

3.1 System Overview

Fig. 1 shows our project's system overview. Two users are in different physical locations and join the developed training drill, which takes place in a virtual space shared through the Internet. Inside the shared virtual space we have synthesized fire evacuation training drill that are generated by using our optimization-based method. Participants are able to extinguish the fires by using an integrated fire extinguisher that will show up on their hands when they enable it. The users can communicate with each other inside the virtual environment freely through Voice over Internet Protocol (VoIP). We placed virtual agents who can respond to participants' voice commands and need to be rescued. The participants' common task is to guide all the agents outside the building.

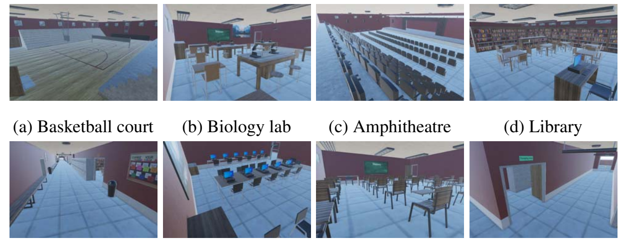
3.2 Environment Representation

We represent the input training environment as an $M \times N$ in size 2D grid ($[c_{1,1}, ..., c_{M,N}]$ denotes the cells of the generated grid; the resolution of the grid is defined by the designer/trainer). Then, we represent each grid cell ($c_{x,y}$) of the grid as either obstacle (T_{obs}), fire (T_{fire}), or empty (T_{empty}) grid cell.

3.3 Virtual Training Environment

We designed a virtual school layout according to specific design and safety regulations[1] and standards in the US [2]. We have created several types of classrooms (standard classroom, library, basketball court, theater, restrooms, lockers, etc.) to convey a complete impression of a school. The average size of a classroom is 12×12 m with a height of 3.75 m to ensure that participants can move around fast and freely while avoiding virtual objects/obstacles (desk, chairs, etc.). Finally, we have decided to add a significant number of exits (six in total) to ensure that users can find accessible exits under different conditions and effects that block some or most of them. Fig. 2 shows screenshots of the designed virtual environment.

[1] https://www.aps.edu/facilities-design-and-construction/design-standards-and-guidelines



(a) Basketball court (b) Biology lab (c) Amphitheatre (d) Library

(e) Corridor (f) Computer lab (g) Lecture room (h) Lockers & Exit

Figure 2: Different parts of the designed virtual environment we used in our prototype application.

The virtual agents can respond to specific voice commands under certain conditions (see Fig. 3). There are six usable commands implemented in the system. Among them, we implemented four commands to instruct the agents to move, including "come here," "follow me," "run," and "crawl." We also included the "stop" and "wait" commands to pause the movement of agents at any time.

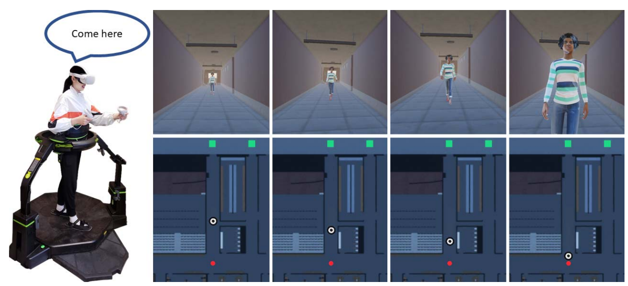


Figure 3: A user commands a virtual agents to "come here" and the agent moves toward the user.

3.4 Authoring Training Drill

We represent our training drill as a composition of several fires $F = [f_1, ..., f_K]$ taking places in a static 3D environment. Their position and size are determined based on our proposed optimization-based method (see Section 4). There are also several trainer-defined virtual agents $A = [a_1, ..., a_B]$ placed in different locations in the virtual environment. The trainer instructs the users of our training drill to rescue the virtual agents by helping them exit the building.

4 Problem Formulation

The design of the evacuation drill d is evaluated by the total cost function $C_{\text{Total}}(d)$:

$$C_{\text{Total}}(d) = w_{\text{length}} C_{\text{length}}(d) + w_{\text{turns}} C_{\text{turns}}(d) + w_{\text{fire}} C_{\text{fire}}(d) + w_{\text{vis}} C_{\text{vis}}(d), \quad (1)$$

where C_{length} encodes the length of the optimal path that the user should follow to fulfil the necessary goals and exit the building; C_{turns} encodes the number of turns in the optimal path; C_{fire} denotes the number of fires that the user should extinguish to fulfil the necessary goals (e.g., access the virtual agents, help virtual agents exit the building); and C_{vis} denotes the visibility conditions of the virtual environment. w_{length}, w_{turns}, w_{fire}, and w_{vis} are the corresponding weights of each cost term, prioritized by importance. We discuss the details for each cost term as follows.

Length Cost. The path synthesized by our system represents how far the user must walk in the training environment to execute the required task. The length cost is used to compare the length of the synthesized path against the user-defined target path length. We present this cost as:

$$C_{\text{length}}(d) = \frac{1}{L_{\text{diag}}}\left|\sum_{\mathcal{G}_i(A)} L(P_i) - \rho_{\text{length}}\right|, \tag{2}$$

where L_{diag} is used as a normalization term representing the diagonal length of the entire virtual environment; ρ_{length} denotes the user-defined path length; P_i is the path between each $\mathcal{G}_i(A)$ sub-group of agents that are in a specific location (e.g., in the basketball court) in the virtual environment requiring rescue, where $\mathcal{G}_i(A) \leqslant A$; and $L(P_i)$ is the distance between the $i-th$ sub-group of agents $\mathcal{G}_i(A)$ and the closest exit in the training environment. To compute the length of the chosen optimal path, we use an improved version of the A* algorithm [35]. For each returned pair of adjacent cells (c_j, c_{j+1}) belonging to the P_i path in the grid, we compute path length $L(P_i)$ by summing the length of each pair of adjacent cells $\mathcal{L}(c_j, c_{j+1})$ from the optimal path as:

$$L(P_i) = \sum_{c_j, c_{j+1}}^{|P_i|-1} \mathcal{L}(c_j, c_{j+1}), \tag{3}$$

where $|P_i|$ denotes the total number of grid cells from the optimal path. Note that the obstacle T_{obs} and fire T_{fire} grid cells are blocked, and the empty grid cell T_{empty} is unblocked. However, during the optimization process, if there is no optimal path, we label the fire grid cell as unblocked, and therefore it can be considered part of the optimal path. Thus, the synthesized path length comes closer to the target path length and makes the training drill more difficult since the user needs to extinguish a fire to access that path properly.

Turn Cost. The turn cost is used to compare the number of turns in the path against a user-defined target number of total turns ρ_{turns}:

$$C_{\text{turns}}(d) = \left|\frac{\sum_{|P|}\mathcal{T}(P_i) - \rho_{\text{turns}}}{\rho_{\text{turns}}}\right|, \tag{4}$$

where $\mathcal{T}(P_i)$ returns the number of turns in the optimal path P_i, and $|P|$ denotes the total number of optimal paths the users should follow to accomplish the task. To calculate $\mathcal{T}(P_i)$, we consider all triads of adjacent grid cells. If these three grid cells do not form a straight line, they are regarded as a turn and, therefore, $\mathcal{T}(P_i)$ returns 1; otherwise, it returns 0.

Fire Cost. Users must extinguish fires to reach virtual agents, access parts of the virtual building, or exit the virtual building. The fire cost compares the number of fires that the user should extinguish against the designer-specified target number of fires ρ_{fire}:

$$C_{\text{fire}}(d) = \frac{1}{U}\left|\sum_{|F|}\Gamma(f_i) - \rho_{\text{fire}}\right|, \tag{5}$$

where $\Gamma(f_i)$ returns 1 if f_i is found to be in the optimal path; otherwise, it returns 0. U is used as a normalization factor representing the upper limit of the number of fires. We set $U = 40$ as the upper limit value for all examples presented in this paper.

Visibility Cost. The user's visibility in the virtual environment is computed by considering the ratio between the area occupied by the fires over the total area of the virtual environment. We compare it against a user-defined target value:

$$C_{\text{vis}}(d) = \left|\frac{\sum_{|F|}\mathcal{A}(f_i)}{\mathcal{A}(e)} - \rho_{\text{vis}}\right|, \tag{6}$$

where $\sum_{|F|}\mathcal{A}(f_i)$ represents the total area occupied by the fires; $\mathcal{A}(e)$ represents the total area of the entire virtual environment; and ρ_{vis} is user-defined target visibility. Note that a high value of $\rho_{\text{vis}} \in [0, 1]$ denotes low visibility and vice versa.

5 Optimization

To assess all possible training outcomes during the optimization process, our system optimizes total cost functions through the reversible-jump Markov chain Monte Carlo (RJMCMC) method [12]. We apply simulated annealing using a Metropolis-Hastings state-search step [7]. We start by defining a Boltzmann-like objective function:

$$f(d) = \exp\left(-\frac{1}{t}C_{\text{Total}}(d)\right), \tag{7}$$

where t encodes the temperature parameter of simulated annealing. During the optimization process, the system proposes a new configuration of the training drill d' by altering the current training drill d using one of the following moves:

- **Adding a fire:** Our system places a randomly sized fire in a randomly chosen position in the virtual environment.
- **Removing an existing fire:** Our system randomly chooses a fire from the virtual environment to remove.
- **Modifying an existing fire:** Our system randomly chooses a fire from the virtual environment and modifies its size and position.

We set the probability of adding a fire as $p_{\text{add}} = .40$, the probability of removing a fire as $p_{\text{remove}} = .20$, and the probability of modifying a fire as $p_{\text{modify}} = .40$. Through these probabilities, our system chooses to add and modify a fire more often than choosing to remove a fire. By applying one of these moves, our system proposes a training drill d' and compares the total cost of the proposed training drill $C_{\text{Total}}(d')$ with the total cost of the current training drill $C_{\text{Total}}(d)$ to determine whether the system accepts the proposed training drill d' or keeps the current training drill d.

To ensure balanced trans-dimensional optimization, we define the probability of each move. Our system computes the probability of adding a fire as:

$$p_{\text{add}}(d'|d) = \min\left(1, \frac{p_{\text{remove}}}{p_{\text{add}}}\frac{U-|d|}{|d'|}\frac{f(d')}{f(d)}\right); \tag{8}$$

computes the probability of removing an existing fire as:

$$p_{\text{remove}}(d'|d) = \min\left(1, \frac{p_{\text{add}}}{p_{\text{remove}}}\frac{|d|}{U-|d'|}\frac{f(d')}{f(d)}\right); \tag{9}$$

and computes the probability of modifying an existing fire as:

$$p_{\text{modify}}(d'|d) = \min\left(1, \frac{f(d')}{f(d)}\right). \tag{10}$$

Based on the above formulation, we set an upper limit on the number of fires during optimization using the variable $U = 40$. Thus, our system synthesizes a virtual environment with fires equal to or less than U.

We also applied simulated annealing to explore our solution space effectively. Simulated annealing allows us to use a temperature parameter t to control the acceptance probability of the proposed training drill d'. If the temperature parameter is high, the system will aggressively explore the whole solution space. If the temperature parameter is low, the optimizer will become more selective. We initialize the temperature parameter as $t = 1.00$ at the beginning of optimization. In each iteration, we multiply the temperature parameter by 0.998. The optimization process terminates when the change in $C_{\text{Total}}(d)$ is less than 5% of the previous 50 iterations.

Unless specified otherwise, we set the weight of the length cost to $w_{\text{length}} = 1.00$, the weight of the turn cost to $w_{\text{turns}} = .40$, the weight of the fire cost to $w_{\text{fire}} = .60$, and the weight of the visibility cost to $w_{\text{vis}} = .40$. Via those weights, our system prioritizes the length of the path and the number of fires the user must extinguish. However, the designer may change the priority by changing the weights.

6 User Study

The user study was conducted between two universities (Purdue and GMU) across states in the US. The two universities were not in the same physical spaces. The intent of our project is to evaluate whether our proposed method can synthesize training drills with different targeted difficulty levels, thus triggering any difference in the collaboration behavior among participants. The methodology of the study is described in the following subsections. Fig. 4 shows example scenes from the synthesized training drill.

Figure 4: Example scenes from the synthesized training drill.

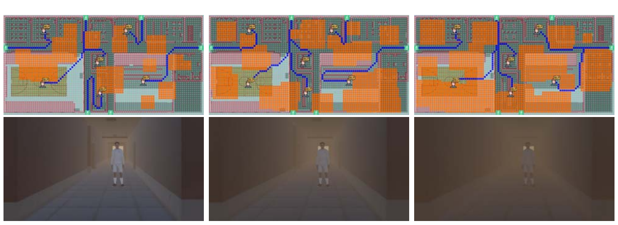
6.1 Participants

We recruited participants in both universities via class announcements and emails. Participants from each university were randomly assigned to a group. Each group was scheduled to attend the study simultaneously at each location. Participants in the same group remotely joined the shared virtual space to experience the synthesized training drills. We collected data from 27 groups (54 volunteers; 34 male and 20 female). The age of the participants were between 17-30 years ($M = 19.96$, $SD = 2.88$). All participants have experienced virtual reality before.

6.2 Conditions

We developed three experimental conditions to determine whether the optimized training drills with differently targeted difficulty would influence the collaboration behaviors among the participants. The experiment followed a within-group study design. We used the Latin squares [33] ordering method to balance the conditions and minimize the carryover effects. Fig. 5 shows the three synthesized training drills used in our experiment. The conditions were as follows:

- **Low Difficulty (LD):** We set the target cost terms as: $\rho_{\text{length}} = 280$, $\rho_{\text{turns}} = 30$, $\rho_{\text{fire}} = 3$, and $\rho_{\text{vis}} = .20$.
- **Medium Difficulty (MD):** We set the cost terms as: $\rho_{\text{length}} = 300$, $\rho_{\text{turns}} = 35$, $\rho_{\text{fire}} = 5$, and $\rho_{\text{vis}} = .50$.
- **High Difficulty (HD):** We set the cost terms as: $\rho_{\text{length}} = 320$, $\rho_{\text{turns}} = 40$, $\rho_{\text{fire}} = 7$, and $\rho_{\text{vis}} = .80$.



(a) Low Difficulty (b) Medium Difficulty (c) High Difficulty

Figure 5: The three experimental conditions we used for our user study. Top: The position and size of fires (orange cells), the optimal paths (blue cells), and the position of the virtual agents. Bottom: The visibility of each training drill.

6.3 Measurements

We collected participants' perceived mutual awareness, mutual assistance, and dependent actions based on the questionnaire developed by Biocca et al. [4]. For each question, we used a 7-point Likert scale. In addition, we collected several in-game measurements to record participants' collaborative behavior. These in-game measurements include the completion time, completion time offset, trajectory length, distance between participants, extinguisher counts, and number of commands.

6.4 Procedure

After we grouped the volunteers, we scheduled each group a specific time slot to attend the study at their corresponding university campus. Once both participants arrived, we first asked them to sign the consent form, which was approved by each university's Institutional Review Board (IRB), if they agreed to participate. Next, the research team collected the demographic information from the participants by asking them to fill out a questionnaire. Then, our research team introduced and helped the participants with the experiment procedures and virtual reality equipment.

Participants first joined the warm-up session to meet in the warm-up scene; integrating a tutorial session improves participants' performance and experience [14]. The warm-up scene was different from the experiment scenes, but all the interaction mechanisms were the same. We instructed them to familiarize themselves with the voice commands and their functionality. Next, the research team informed them how to use the fire extinguisher and enable the minimap (see Fig. 6a), and at the same time, they became familiar with the Virtuix Omni treadmill. Once participants finished the warm-up session and agreed to start the experiment, the research team helped them join the experiment's scene. In Fig. 6b, we show two users trying to open a path using the fire extinguisher. The warm-up session took no more than five minutes, and each experiment session lasted about 10 minutes (no participant spent more than one hour to complete the entire study). We informed participants they were allowed to give up the study; however, no participant quit.

(a) (b)

Figure 6: (a) Users can enable a minimap. The minimap provided information on players' position, the position of the virtual agents, the exits, and the commands they could use. (b) Two users collaborate in the shared virtual environment to open a path to escape the building.

6.5 Setup and Implementation Details

We used Unity Game Engine 2020.3.20f1 to develop the application. We also used a Dell Alienware Aurora R7 desktop computer (Intel Core i7, NVIDIA GeForce RTX 2080, 32GB RAM) in each university to run the application. We used Unity's Photon asset to implement the network frame to allow participants to communicate and collaborate in a shared virtual space. The optimization process for authoring each training drill did not exceed 30 seconds. We used the Virtuix Omni treadmill to allow participants to move around in the virtual environment and Occulus Quest 2 as a VR headset. Lastly, we used the `KeywordRecognizer` class provided by Microsoft and integrated it into the UnityEngine library for voice recognition.

6.6 Results

We used one-way repeated measures analysis of variance (ANOVA) to analyze the data collected from three experimental conditions (LD, MD, and HD). We assessed the individual differences using post-hoc Bonferroni corrected estimates if the ANOVA was statistically significant. We provide the descriptive statistics in the supplementary materials file.

6.6.1 Objective Data

The analysis revealed a statistically significant result for the **completion time** measurement across the three examined conditions ($\Lambda = .413$, $F[2,25] = 17.791$, $p = .000$, $\eta_p^2 = .587$). The post-hoc pairwise comparison showed that the completion time during the LD condition was significantly lower than that for the MD ($p = .030$) and HD ($p = .000$) conditions. Moreover, the completion time was significantly lower for the MD condition than the HD condition ($p = .012$). We also found a statistically significant result for the **extinguisher count** measurement ($\Lambda = .381$, $F[2,52] = 42.179$, $p = .000$, $\eta_p^2 = .619$). The post-hoc pairwise comparison revealed that our participants used the virtual extinguisher less often in the LD condition than the MD ($p = .000$) and HD ($p = .000$) conditions; moreover, the participants used the virtual extinguisher less often during the MD condition than the HD condition ($p = .019$). However, the statistical analysis did not reveal significant results for the **completion time offset** ($\Lambda = .966$, $F[2,25] = .441$, $p = .649$, $\eta_p^2 = .034$), **trajectory length** ($\Lambda = .942$, $F[2,52] = 1.592$, $p = .213$, $\eta_p^2 = .058$), **distance between participants** ($\Lambda = .883$, $F[2,25] = 1.663$, $p = .210$, $\eta_p^2 = .117$), and **number of commands** ($\Lambda = .962$, $F[2,52] = 1.033$, $p = .363$, $\eta_p^2 = .038$). We provide the descriptive statistics in the supplementary materials file.

6.6.2 Subjective Self-reported Data

The **mutual awareness** measurement was statistically significant ($\Lambda = .618$, $F[2,52] = 16.062$, $p = .000$, $\eta_p^2 = .382$) across the three examined conditions. The post-hoc pairwise comparison showed that mutual awareness was significantly lower during the LD condition than the MD ($p = .000$) and HD ($p = .000$) conditions. Similarly, **mutual assistance** was statistically significant ($\Lambda = .593$, $F[2,52] = 17.877$, $p = .000$, $\eta_p^2 = .407$). The post-hoc pairwise comparison revealed that mutual assistance was significantly lower during the LD condition than the MD ($p = .034$) and HD ($p = .000$) conditions, and the MD condition was significantly lower than the HD condition ($p = .001$). The **dependent actions** measurement was also statistically significant across the three conditions ($\Lambda = .286$, $F[2,52] = 64.943$, $p = .000$, $\eta_p^2 = .717$). The post-hoc pairwise comparison showed that dependent actions were rated significantly lower during the LD condition than the MD ($p = .000$) and HD ($p = .000$) conditions, and the MD condition was rated significantly lower than the HD condition ($p = .000$).

6.7 Discussion

The collected objective data, and more specifically the **completion time** and **extinguisher count** measurements, revealed that our method can automatically synthesize training drills that have different difficulty levels for executing them. These findings prove that it is possible to synthesize fire evacuation training drills in which the trainer/designer can specify the parameters, such as the path length, number of turns in the optimal paths, number of fires, environment visibility, and the system can synthesize variations of the training drill without impacting the overall objective of that drill. However, the **trajectory length** measurement was not statistically significant across the three examined conditions. Considering that our participants walked the same trajectory lengths across the three conditions, the completion time proves that they needed more time to complete a more difficult training drill in comparison to the MD or LD training drills, in which they extinguish fewer fires and had higher visibility. If we also consider the **number of commands** measurement we could say that our participants tried to instruct the virtual agents in roughly the same way across the three conditions. Thus, we can say that the virtual fires (due to **completion time** and **extinguisher count**) impacted our participants' behavior in executing the tasks, but not the virtual agents. Consequently, we argue that our method can synthesize training drills based on the difficulty entailed in executing them.

In contrast, the other measurements did not differ across the three experimental conditions. Specifically, an interesting observation was made for the **completion time offset** and the **distance between participants** measurements. In both measurements, although the completion time offset and the distance between participants decreased from the LD condition to the MD condition and from the MD condition to the HD condition, the decreases were not statistically significant. However, by looking at the mean values for the completion time offset measurement, it is evident that the time offset is close to 30 seconds for all three conditions. A similar observation can be made for the distance between the participants: their mean distance is sufficient across the three conditions, which indicates that they were in different locations in the building during the training drills. These findings suggest that although the participants were in the same shared space, they chose their strategies and acted independently. Such independent activity has been identified by Tang et al. [28] as the "same problem, different area" style of coupling between two people. Therefore, we think that our two participants preferred to utilize a collaborative behavior that could help them execute the given task in a way that was more optimal for them.

The **mutual awareness** measurement indicated that the participants were aware of each other during the training drill. It seems that the difficulty of the training drill impacted their awareness of one another. Therefore, the participants felt they were not alone while executing the given task in the virtual environment The **mutual assistance** and **dependent actions** measurements revealed that, as the difficulty level of the training drill increased, the mutual assistant of each participant (the degree to which each person needed to help the other person) and their perceived dependence on the other participant increased. These findings indicate that the participants felt the pressure of the training drill, and they tried to assist the other participant by creating a strategy that would help them execute the given task and assist the other person.

Overall, by combining both the objective and self-reported measurements, we can say that, though our participants planned their strategy independently of each other, they were always aware of the other individual in the shared virtual environment, and given their awareness, they planned their strategy to help not only themselves but also the other participant. It looks as if this kind of planning is common in games [1] where players on the same team work together to accomplish a given task. Our results showed that, though the two participants were in separate locations, being in a shared virtual space and sharing the same goals and tasks made them choose individual strategies that benefited themselves and the team; therefore, establishing a collaborative culture.

6.8 Limitations

Our study had some limitations. First, our participants were not exposed to real-world evaluations. Therefore, we cannot firmly conclude that the training platform and its performance are effective in real-world emergency evacuation scenarios. Second, due to the hardware limitations (we used an Omni treadmill), long-time locomotive tasks will result in the users needing to exert physical effort and experience fatigue [23], which could potentially decrease their motivation. Third, our optimization-based approach only considered four design decisions to synthesize the training drill. We think additional cost terms could be considered, such as those related to specific training objectives.

7 Conclusion

In this paper, we introduced a method to synthesize training drills for fire evacuation scenarios. Due to the proposed optimization-based formulation, a designer/trainer can easily define the target objectives for each cost term. Our system automatically synthesizes the training scenario where participants encounter the specified difficulty of executing a task. Thus a designer/trainer could easily generate several variations of a training drill, allowing trainees to experience them and get prepared for potential real-world situations.

Acknowledgements

This project is partially supported by an NSF CAREER grant (award #: 1942531) and an NSF FW-HTF-R grant (award #: 2128867).

References

[1] S. A. Alharthi, R. C. Torres, A. S. Khalaf, Z. O. Toups, I. Dolgov, and L. E. Nacke. Investigating the impact of annotation interfaces on player performance in distributed multiplayer games. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2018.

[2] L. Baker. A history of school design and its indoor environmental standards, 1900 to today. *National Clearinghouse for Educational Facilities*, 2012.

[3] E. Z. Barsom, M. Graafland, and M. P. Schijven. Systematic review on the effectiveness of augmented reality applications in medical training. *Surgical endoscopy*, 30(10):4174–4183, 2016.

[4] F. Biocca, C. Harms, and J. Gregg. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *4th annual international workshop on presence, Philadelphia, PA*, pp. 1–9, 2001.

[5] J. C. Burguillo. Using game theory and competition-based learning to stimulate student motivation and performance. *Computers & education*, 55(2):566–575, 2010.

[6] F.-K. Chiang, X. Shang, and L. Qiao. Augmented reality in vocational training: A systematic review of research and applications. *Computers in Human Behavior*, 129:107125, 2022.

[7] S. Chib and E. Greenberg. Understanding the metropolis-hastings algorithm. *The american statistician*, 49(4):327–335, 1995.

[8] P. Dillenbourg. *Collaborative learning: Cognitive and computational approaches. advances in learning and instruction series.* ERIC, 1999.

[9] C. Elvezio, M. Sukan, O. Oda, S. Feiner, and B. Tversky. Remote collaboration in ar and vr using virtual replicas. In *ACM SIGGRAPH 2017 VR Village*, pp. 1–2. 2017.

[10] L. Freina and M. Ott. A literature review on immersive virtual reality in education: state of the art and perspectives. In *The international scientific conference elearning and software for education*, vol. 1, pp. 10–1007, 2015.

[11] R. Gajda. Utilizing collaboration theory to evaluate strategic alliances. *American journal of evaluation*, 25(1):65–77, 2004.

[12] P. J. Green. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82(4):711–732, 1995.

[13] S. W. Greenwald, A. Kulik, A. Kunert, S. Beck, B. Fröhlich, S. Cobb, S. Parsons, N. Newbutt, C. Gouveia, C. Cook, et al. Technology and applications for collaborative learning in virtual reality. Philadelphia, PA: International Society of the Learning Sciences., 2017.

[14] D. Kao, A. J. Magana, and C. Mousas. Evaluating tutorial-based instructions for controllers in virtual reality games. *Proceedings of the ACM on Human-Computer Interaction*, 5(CHI PLAY):1–28, 2021.

[15] A. D. Kaplan, J. Cruit, M. Endsley, S. M. Beers, B. D. Sawyer, and P. Hancock. The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis. *Human factors*, 63(4):706–726, 2021.

[16] G. Kurillo, R. Bajcsy, K. Nahrsted, and O. Kreylos. Immersive 3d environment for remote collaboration and training of physical activities. In *2008 IEEE Virtual Reality Conference*, pp. 269–270. IEEE, 2008.

[17] E. R. Lai. Collaboration: A literature review. *Pearson Publisher. Retrieved November*, 11:2016, 2011.

[18] K. Lee. Augmented reality in education and training. *TechTrends*, 56(2):13–21, 2012.

[19] C. Li, W. Liang, C. Quigley, Y. Zhao, and L.-F. Yu. Earthquake safety training through virtual drills. *IEEE transactions on visualization and computer graphics*, 23(4):1275–1284, 2017.

[20] H. Liu, Z. Wang, C. Mousas, and D. Kao. Virtual reality racket sports: Virtual drills for exercise and training. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 566–576. IEEE, 2020.

[21] A. Luxenburger, A. Prange, M. M. Moniri, and D. Sonntag. Medicalvr: Towards medical remote collaboration using virtual reality. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pp. 321–324, 2016.

[22] B. R. Martins, J. A. Jorge, and E. R. Zorzal. Towards augmented reality for corporate training. *Interactive Learning Environments*, pp. 1–19, 2021.

[23] C. Mousas, D. Kao, A. Koilias, and B. Rekabdar. Evaluating virtual reality locomotion interfaces on collision avoidance task with a virtual character. *The Visual Computer*, 37(9):2823–2839, 2021.

[24] T. Piumsomboon, G. A. Lee, and M. Billinghurst. Snow dome: A multi-scale interaction in mixed reality remote collaboration. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–4, 2018.

[25] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst. Mini-me: An adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pp. 1–13, 2018.

[26] J. Roschelle and S. D. Teasley. The construction of shared knowledge in collaborative problem solving. In *Computer supported collaborative learning*, pp. 69–97. Springer, 1995.

[27] D. D. Ruikar, R. Hegadi, and K. C. Santosh. A systematic review on orthopedic simulators for psycho-motor skill and surgical procedure training. *Journal of Medical Systems*, 42:1–21, 2018.

[28] A. Tang, M. Tory, B. Po, P. Neumann, and S. Carpendale. Collaborative coupling over tabletop displays. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 1181–1190, 2006.

[29] S. Tea, K. Panuwatwanich, R. Ruthankoon, and M. Kaewmoracharoen. Multiuser immersive virtual reality application for real-time remote collaboration to enhance design review process in the social distancing era. *Journal of Engineering, Design and Technology*, 2021.

[30] R. M. Viglialoro, S. Condino, G. Turini, M. Carbone, V. Ferrari, and M. Gesi. Augmented reality, mixed reality, and hybrid approach in healthcare simulation: a systematic review. *Applied Sciences*, 11(5):2338, 2021.

[31] A. Villanueva, Z. Zhu, Z. Liu, F. Wang, S. Chidambaram, and K. Ramani. Colabar: A toolkit for remote collaboration in tangible augmented reality laboratories. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW1), apr 2022.

[32] S. Webel, U. Bockholt, T. Engelke, N. Gavish, M. Olbrich, and C. Preusche. An augmented reality training platform for assembly and maintenance skills. *Robotics and autonomous systems*, 61(4):398–403, 2013.

[33] E. J. Williams. Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry*, 2(2):149–168, 1949.

[34] B. Xie, H. Liu, R. Alghofaili, Y. Zhang, Y. Jiang, F. D. Lobo, C. Li, W. Li, H. Huang, M. Akdere, C. Mousas, and L.-F. Yu. A review on virtual reality skill training applications. *Frontiers in Virtual Reality*, 2:645153, 2021.

[35] J. Yao, C. Lin, X. Xie, A. J. Wang, and C.-C. Hung. Path planning for virtual human motion using improved a* star algorithm. In *2010 Seventh international conference on information technology: new generations*, pp. 1154–1158. IEEE, 2010.