Safety Science 197 (2026) 107111

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

Investigating exit choices in fire evacuation using multi-user virtual reality experiments

Song Lu [a], Jieyu Chen [a], Nan Li [a,b,*], Zhenan Feng [c], Milad Haghani [d], Ruggiero Lovreglio [c]

[a] Department of Construction Management, Tsinghua University, Beijing 100084, China
[b] Hang Lung Center for Real Estate, Tsinghua University, Beijing 100084, China
[c] School of Built Environment, Massey University, Auckland 0632, New Zealand
[d] Centre for Spatial Data Infrastructures and Land Administration, Faculty of Engineering and Information Technology, The University of Melbourne, Melbourne 3000, Australia

ARTICLE INFO

Keywords:
Evacuation
Exit choice
Group decision-making
Multi-user
Virtual reality
Discrete choice model

ABSTRACT

Understanding exit choice behavior during evacuations is critical for enhancing safety and efficiency in real-life emergencies. Individual, social and environmental factors can significantly influence group decision-making during evacuations. While a substantial amount of research has focused on individual exit choice behaviors, very few studies have investigated how group decision-making is affected by various factors during emergencies, especially when group splitting occurs. This study investigates the impacts of individual, social, and environmental factors on both individual and group exit choice decision-making, using a multi-user virtual reality (VR) setup that simulates fire evacuation scenarios. We conducted experiments with 127 participants, organized in groups ranging from 1 to 4 group members. We observed a clear tendency for 2-, 3-, and 4-person groups to split into smaller units during evacuations, with 2-person subgroups being the most common and larger groups showing a higher likelihood of splitting. Then, a discrete choice model was used to analyze the effects of various social, individual and environmental factors on the participants' exit choices. The results showed that participants were not only influenced by the factors significant in individual decision-making, such as distance to exits, familiarity with exits, and smoke, but also by the choices of other group members, demonstrating a strong social influence. Additionally, the impact of smoke was more pronounced in subgroup settings. This research highlights the significant differences between decision-making at the individual and subgroup levels during fire evacuations, providing insights for developing more sophisticated evacuation simulation models and building safety management protocols.

1. Introduction

Fire evacuation is a crucial aspect of public safety, where the decision-making process of choosing an exit can greatly impact evacuation efficiency and safety (Haghani et al., 2020; Lovreglio et al., 2014). Understanding human exit choice behavior during emergencies is fundamental to improving safety protocols and infrastructure in built environments (Chen et al., 2021; Kuligowski, 2016). Previous studies have explored various factors influencing evacuees' exit choice, including familiarity, social influence, building geometry, and harmful conditions within the environment (e.g., smoke) (Kinateder et al., 2018; Lovreglio et al., 2022; Ren-Yong & Hai-Jun 2010).

Many of these studies have utilized physical and virtual experiments. For instance, Benthorn & Frantzich (1999) observed that evacuees prefer familiar exits, even when they are farther than the nearest exit, although emergency exits become more attractive if visible and accessible. Similar findings in real fire events (Fridolf et al., 2013; Grosshandler et al., 2005) and evacuation drills (Rahouti et al., 2021) highlight this preference for familiar exits. Kinateder et al. (2018) further confirmed this in virtual experiments, showing that evacuees disproportionately favor familiar exits, especially when influenced by neighbors' behavior. Bode & Codling (2013) investigated the role of social influence in virtual crowd evacuations, revealing that evacuees are more likely to follow the crowd, which can lead to less optimal exit choices. Kinateder & Warren (2021) explored how social influence and building geometry interact, showing that in smaller crowds, evacuees

* Corresponding author.
E-mail address: nanli@tsinghua.edu.cn (N. Li).

https://doi.org/10.1016/j.ssci.2026.107111
Received 1 August 2025; Received in revised form 8 January 2026; Accepted 9 January 2026
Available online 3 February 2026

are more likely to follow the majority, while in larger crowds, they avoid congested exits if the doors are narrow but are less concerned when the doors are wide. Additionally, virtual experiment studies have shown that environmental conditions like smoke significantly reduce the likelihood of evacuees choosing certain exits (Gao et al., 2022; Lovreglio et al., 2022). These studies have provided insights into individual exit choice behaviors during evacuations. For a more comprehensive review on exit choice experiments, readers are invited to the detailed reviews by Feng et al. (2021b) and Haghani & Sarvi (2018).

Despite the contributions of these individual-focused studies, three explicit gaps and limitations persist in the current literature, which hinder the development of comprehensive group evacuation theories and practical safety strategies. First, most existing group evacuation research prioritizes "evacuation dynamics" (e.g., evacuation time, movement speed) over "exit choice decisions", which is the foundational step that directly determines whether evacuees can reach safety efficiently. For example, Bode et al. (2015) and Lu et al. (2017) only analyzed how social groups affect evacuation time in crowds, while Zhang et al. (2018) focused on movement speed differences between individuals and small groups in stairwells; none of these studies addressed how groups select exits, especially when their cohesion breaks. Second, the limited number of studies examining group exit choices (e.g., Haghani et al., 2019) enforce artificial group cohesion, requiring all members to evacuate together and only simulating scenarios with strong social ties (e.g., families). This design fails to reflect real emergencies, where groups with weaker social ties (e.g., casual friends, colleagues) often split into smaller units (Cornwell & Ho, 2022; Felicianiet al., 2023), leaving the exit choice mechanisms of subgroups and individuals post-splitting entirely unstudied. Third, discrete choice models widely used to quantify individual exit choice factors (Lovreglio et al., 2022; Duives & Mahmassani, 2012) have not been applied to compare the trade-off processes between subgroups and individuals after a split. This lack of quantitative analysis means we cannot determine how key factors (e.g., smoke, social influence) differentially impact these two decision-making entities, a critical gap for calibrating accurate evacuation simulation models.

Based on the three critical gaps identified earlier: (1) existing group evacuation research focuses on "evacuation dynamics" (e.g., time, speed) rather than "exit choice of subgroups/individuals post-splitting", (2) artificial mandatory group cohesion ignores natural splitting of weak-tie groups, and (3) discrete choice models have not been used to compare trade-off processes between subgroups and individuals, this study aims to advance the field not by reintroducing the phenomenon of group splitting (a known occurrence), but by directly addressing these unmet needs through targeted, method-driven design. In doing so, this work makes three interrelated contributions:

To tackle the first gap (missing data on post-splitting exit choice), we shift the research focus from group movement dynamics to the core decision of exit selection, collecting empirical data on how subgroups and individuals choose exits after splitting—data that has been absent in prior work. For the second gap (lack of natural weak-tie group simulation), we abandon the constraint of forced cohesion; instead, we simulate realistic weak-tie groups (e.g., acquaintances, colleagues, as noted in Wu & Zheng, 2024; Xia et al., 2024) by encouraging pre-experiment socialization among 140 participants (randomly assigned to groups of 1–4 people) to establish weak social ties, allowing natural splitting behavior to emerge during evacuations. To fill the third gap (unexamined trade-off processes via modeling), we extend the application of discrete choice models by building on and modifying an existing VR environment (Lovreglio et al., 2022) to incorporate key influencing factors (familiarity, crowding, proximity, and smoke-impacted visibility) and to quantitatively analyze and compare how these factors differentially impact the exit choices of individuals and subgroups. This integrated approach not only captures the dynamic decision-making of split groups but also provides actionable parameters for evacuation model optimization, ultimately filling the critical gap between "documenting group splitting" and "understanding the underlying split-group exit choice mechanisms."

The rest of the paper is organized as follows: Section 2 provides detailed descriptions of the methodologies, including the construction of a multi-user VR evacuation environment and the theoretical framework of the discrete choice model. Section 3 presents the results in terms of the group splitting pattern, exit choice model, and participants' responses to subjective questionnaires. Finally, Section 4 discusses and concludes the paper, summarizing key findings and outlining directions for future research.

2. Methodology

This section outlines the VR experimental design (Section 2.1), the multi-user VR system (Section 2.2), the experimental procedure (Section 2.3), the participants (Section 2.4), the measurements (Section 2.5), and the discrete choice modeling (Section 2.6).

2.1. VR experimental design

This section details the design of the VR experiment, including the virtual environment setup, exit characteristic configurations considering key factors influencing participants' decisions, and scenario design. The virtual environment featured a waiting room and a meeting room. Participants began VR experience in the waiting room at a designated Starting Point and moved toward the Initial Point in the meeting room, marked by a red square in Fig. 1. As depicted in Fig. 1, the meeting room had a few pieces of furniture and three exits labeled A, B, and C. The positions of exits B and C varied with changing scenarios, representing different distances participants had to cover to reach the doors during an evacuation. These variations are detailed in Table 1. Above each exit, signs clearly marked them as available evacuation paths. In the experiment, each available exit was designed with distinct characteristics in terms of environmental, social, and individual factors. The environmental aspect was highlighted by the varying distances of the exits from the participant's initial point, coupled with the presence or absence of smoke emerging near some exits. In the social aspect, the choice of exits was influenced by the decisions of non-player characters (NPCs) chosen specific exits and group members who opted for these exits prior to the participants' decision-making. In the individual aspect people are more likely to choose familiar exits during an evacuation (Kinateder et al., 2018). In our experiment, we designated Exit A as the familiar exit for the participants. As illustrated in Fig. 1, Exit A, strategically positioned along the participants' path from the Starting point to the Initial Point, was presumed to be more familiar to participants compared to other exits.

In summary, the participants' choice of exits could be influenced by five variables in the experimental settings: NP (NPCs using the exit), DIST (distance to the exit), SMOKE (presence of smoke), FAM (familiarity with exit), and GM (group members' decisions). These variables fall into different aspects: DIST and SMOKE are environmental factors; FAM is an individual factor; NP and GM are social factors. Except for the GM factor, which was coded based on the participants' behavior during the experiment, all other variables were predetermined during the experimental design, as shown in Table 1. These variables are summarized below.

NP: This variable captures the social aspect of evacuation, varying in levels from 0 to 10 evacuees NPCs per exit. This range aligns with our goal of maintaining crowd density below the critical threshold of 0.54 persons/m^2 as per SFPE guidelines (Gwynne & Rosenbaum, 2016), ensuring realistic evacuation dynamics without overcrowding impacts.

DIST: This variable measures the participant's proximity to each exit. Exit A is consistently at a 6 m distance from the initial point. Exits B and C, however, shift in position, with Exit B being either 3.6 or 5.6 m away, and Exit C at 3 or 4.6 m, as depicted in Fig. 1.

SMOKE: This variable adds a layer of complexity to the decision-

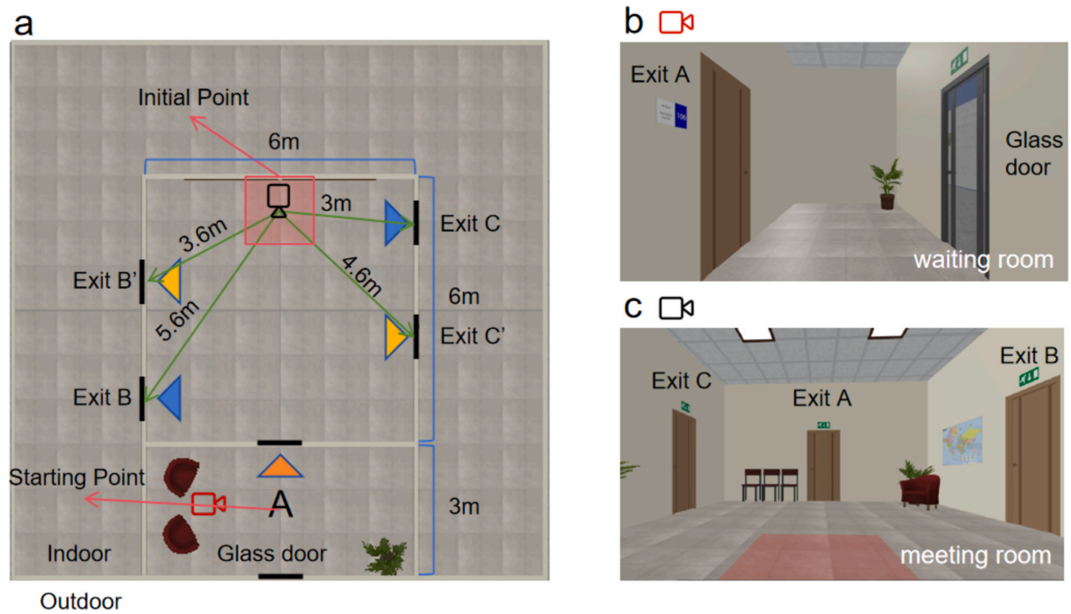

Fig. 1. Virtual environment: (a) geometry of the virtual environment, (b) waiting room view (Starting Point), (c) meeting room view (Initial Point).

Table 1
Definition of all independent variables.

Variable	Levels (values)		
	Exit A	Exit B	Exit C
NPC	4 (0,1,5,10)	4 (0,1,5,10)	4 (0,1,5,10)
DIST	1 (6.0 m)	2 (3.6 m, 5.6 m)	2(3.0 m, 4.6 m)
SMOKE	2 (1, 0)	2 (1, 0)	2 (1, 0)
FAM	1 (1)	1 (0)	1 (0)

Note: The values of each variable are in the parentheses.

making process, with smoke being present (1) or absent (0) at the exits, influencing the perceived safety and navigability.

FAM: In line with our experimental design, familiarity is specific to Exit A, given that participants used it before evacuation. Exit A is deemed as familiar (1), while other exits are marked as unfamiliar (0).

Our scenario design is based on previous single-person experiments. With four variables, three exits, and their various dimensions, our study initially has 2048 possible scenarios (calculated as 4ˆ3 NP × 2ˆ2 DIST × 2ˆ3 SMOKE × 1ˆ3 FAM). By using the Ngene software, we applied an efficient design approach to select the eight most informative scenarios for our experiments (Institute of Transport and Logistics Studies, 2007; SÁndor and Wedel, 2001). These scenarios are detailed in Table 2.

In the experiment, four group sizes were tested: 1, 2, 3, and 4 persons. Participants belonging to the same group were group members, who were distinguished from NPCs by wearing blue clothing in the virtual environment. As a variable reflecting the social factor of group members, GM is defined as follows:

GM: This represents the number of group members who evacuated through certain exit before the participant in an evacuation trial. For each exit, GM reflects the influence of group members on the individual or subgroup's decision, ranging from 0 to 3.

To eliminate ambiguity in data collection and analysis, core terms related to group dynamics are explicitly defined based on pedestrian behavior research conventions (Wu & Zheng, 2024; Feliciani et al., 2023) and the specific experimental context:

Individuals post-splitting: Participants who act independently after group splitting, defined by two criteria: (1) being separated from all other original group members by an Euclidean distance >1.5 m (a threshold reflecting "non-cohesion" in pedestrian group studies); (2) maintaining this separation with any other participant.

Subgroups: Small cohesive units formed after group splitting, meeting three mutually inclusive conditions: (1) Minimum size: ≥2 participants (consistent with the most prevalent unit observed in pilot experiments and prior literature). (2) Spatial proximity: For any participant, the distance to the nearest participant within the same unit does no exceeed 1.5 m (to reflect intentional cohesion).

2.2. Multi-user VR system

Our experiments were based on a multi-user VR system developed by the Urban Resilience Lab at the Department of Construction Management, Tsinghua University. As shown in Fig. 2, this system utilizes Photon PUN2 network architecture, which is a client–server framework supporting cross-platform VR experiences and multi-user access (Photon

Table 2
Eight scenarios selected by efficient design approach using Ngene.

Scenario	Exit A			Exit B			Exit C		
	NP	DIST	SMOKE	NP	DIST	SMOKE	NP	DIST	SMOKE
1	0	6 m	0	10	3.6 m	1	5	4.6 m	1
2	5	6 m	1	0	3.6 m	1	10	3.0 m	0
3	1	6 m	1	1	3.6 m	0	10	3.0 m	0
4	10	6 m	0	0	3.6 m	0	1	4.6 m	1
5	10	6 m	0	1	3.6 m	1	0	4.6 m	0
6	5	6 m	1	10	3.6 m	0	0	3.0 m	1
7	1	6 m	1	5	3.6 m	0	1	3.0 m	0
8	0	6 m	0	5	3.6 m	1	5	4.6 m	1

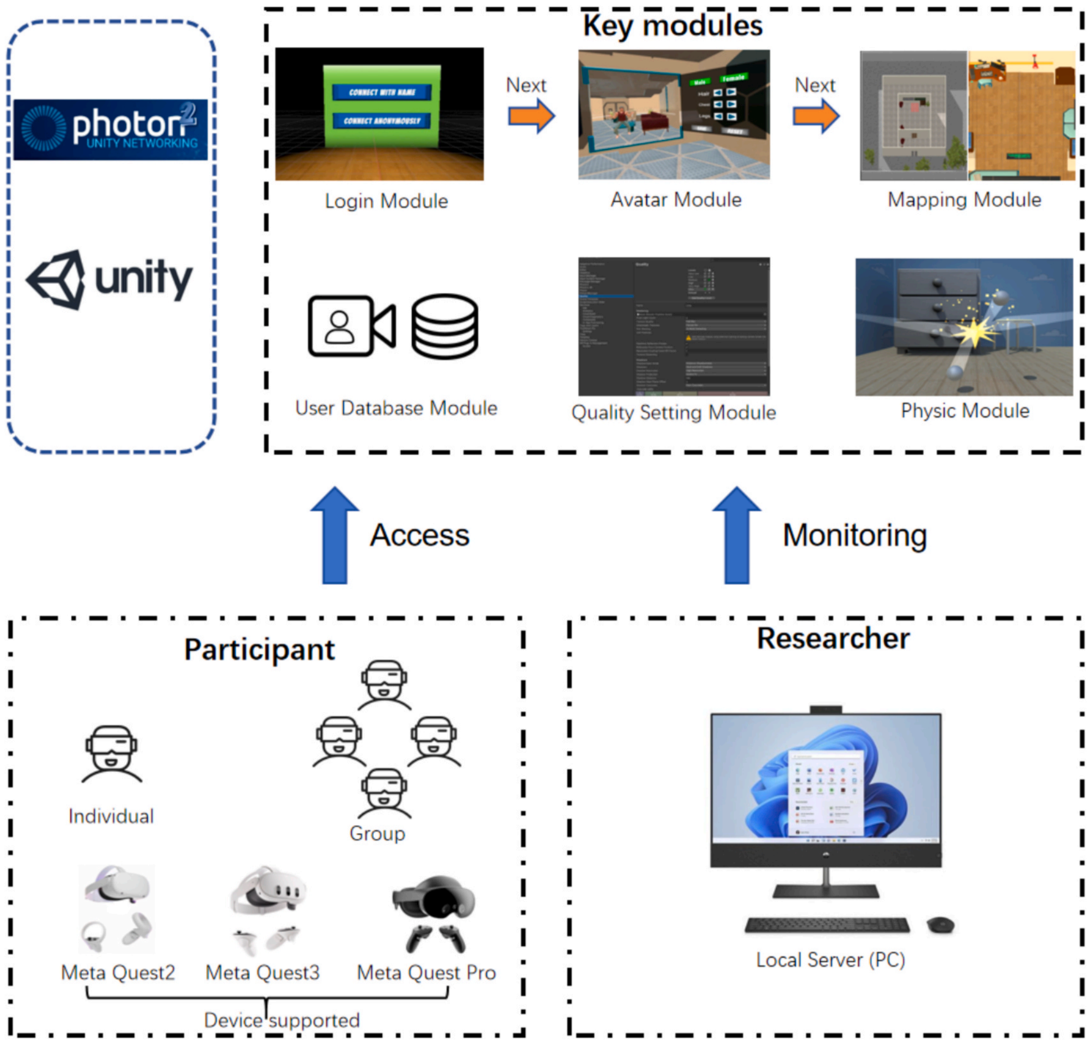

Fig. 2. Architecture of the multi-user VR system for evacuation behavior research.

Pun2 Engine, 2018). Built upon Unity version 2021.3.16f1 (Unity, 2022), the system includes six modules: a Login Module, a Mapping Module, an Avatar Module, a User Database Module, Quality Setting Module, and a Physics Module. The Login Module handles user logins and authorization, linked directly to the User Database. The Mapping Module manages virtual maps within the VR environment and features room creation, customization, and scene transitions. The Mapping Module manages two maps created in Unity, designed for familiarization and experimental stages respectively. The Avatar Module, developed with the Unity packages Final IK and UMA 2, provides customizable virtual avatars and supports full-body movement animation based on controller and headset positioning. Final IK is a key component for avatar movement control, while UMA 2 serves as a comprehensive avatar skinning library. This module also facilitates in-world voice communication by integrating features such as voice storage, stereo sound, and noise suppression, enabling participants to interact audibly within the VR environment. The User Database Module stores user-related data, including login credentials, voice, and tracking data. The Quality Setting Module allows users to adjust graphics and performance to enhance their experience, including resolution, texture details, shadows, and anti-aliasing. Lastly, the Physics Module simulates physical laws, including collision handling, to enhance user immersion. Overall, the system has the necessary functions to provide an immersive, customizable multi-user VR experience to participants and collect essential behavioral data from them.

During our experiments, we utilized a local server equipped with Windows 11, an Intel Core i7-13700K processor, NVIDIA GeForce RTX 4090 graphics card, and 64 GB RAM. VR experiences were provided using Meta Quest series headsets and controllers. Movement was controlled via the joysticks on the VR controllers, while helmet rotation was tracked to synchronize the participants' direction in the VR environment. All behavioral data, including participants' location, orientation, trajectory, and choice of exits, were automatically recorded by the system. Additionally, we utilized EV capture, a screen recording tool, to document the entire experimental process on the server side. This allowed us to monitor participants' movements and orientations in real-time from a top view, facilitating the analysis of group behaviors among participants.

2.3. Experimental procedure

The ethical approval for the research project was granted by the Tsinghua University Ethics Committee on Humanities, Social Sciences, and Engineering, under approval number THU-09-2023-04. As shown in Fig. 3, prior to the experiment, participants were randomly assigned by the experiment administrators to one of the experimental configurations (group sizes of 1, 2, 3, or 4). For group sizes of 2 to 4, participants arrived and participated in the experiment as a group, rather than individually. Upon their arrival at the testing area, participants were presented with information sheets about the study and signed consent forms to confirm

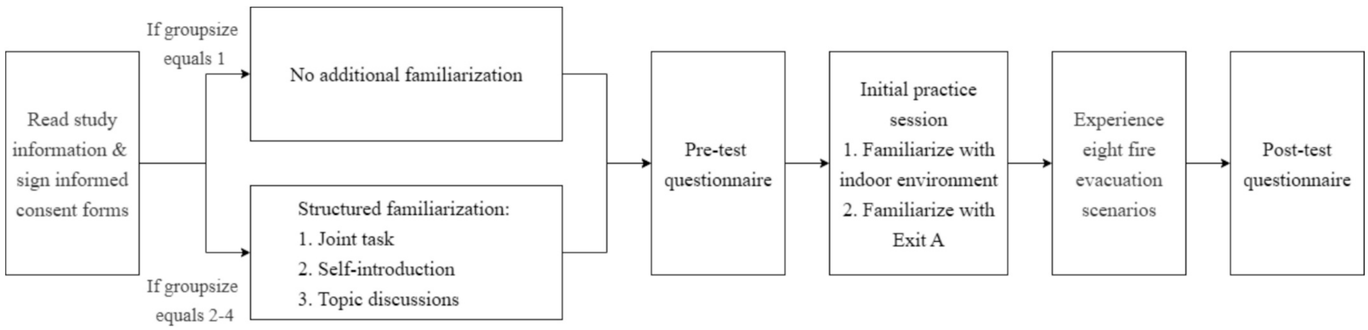

Fig. 3. Experiment procedure diagram.

their voluntary participation. Subsequently, the experimenter equipped the participants with VR headsets, after which they entered a virtual room together. In this room, participants familiarized themselves with the VR equipment. To mitigate potential confounding effects from varying levels of VR proficiency (as reported in Table 3), all participants were required to complete a comprehensive operational training and practice session. This session was designed to minimize the impact of prior operational skill on the experimental outcomes, ensuring all participants had achieved a baseline level of competence with the navigation tasks (e.g., walking, turning). Following the familiarization with the VR equipment, participants assigned to the single-participant configuration (group size 1) did not undergo any additional familiarization process, as they participated individually. For groups of 2 to 4 participants, a structured familiarization process was conducted to encourage socialization and enhance group cohesion. Specifically, they first engaged in a common task within a designated familiarization environment, where they explored the virtual space together and familiarized themselves with their peers' virtual avatars. Such joint activity has been shown to effectively promote cooperation and interaction among group members and enhance group cohesion (Moreland & Levine, 2002). Second, each participant gave a brief self-introduction, including their name, hometown, and hobbies, which helped them quickly get to know each other and establish initial familiarity (Mcgrath, 1991). Third, participants engaged in topic discussions where they shared their views on three topics, including music, movies, and sports. Such discussions are known to further bond group members by fostering interaction (Fisher, 1970). In the experiment, we did not explicitly tell the participants to evacuate as a team or individual, so as not to affect their decision-making. After this familiarization phase, participants were asked to complete the pretest questionnaire, which collected information on their demographics, previous experience (e.g., frequency of practicing in fire drills, frequency of playing video games, and experience with VR) and level of familiarity with group members. After completing the questionnaire, participants donned the VR headsets.

Table 3
The demographics of the participants.

	Group size 1	Group size 2	Group size 3	Group size 4
Group numbers	15	10	12	14
Gender				
Female	6 (40.0%)	10 (50.0%)	22 (61.1%)	26 (46.4%)
Male	9 (60.0%)	10 (50.0%)	14 (38.9%)	30 (53.6%)
Age	23.6	23.8	23.4	22.7
Frequency of practicing in a fire drill				
Never	0 (0.0%)	0 (0.0%)	1 (2.8%)	3 (5.4%)
Less than once a year	6 (40.0%)	11 (55.0%)	22 (61.1%)	28 (50.0%)
Once a year	5 (33.3%)	5 (25.0%)	4 (11.1%)	10 (17.9%)
Twice a year	4 (26.7%)	3 (15.0%)	4 (11.1%)	5 (8.9%)
More than twice a year	0 (0.0%)	1 (5.0%)	2 (5.6%)	7 (12.5%)
Unsure	0 (0.0%)	0 (0.0%)	1 (2.8%)	3 (5.4%)
Frequency of playing video games				
Everyday	5 (33.3%)	6 (28.6%)	12 (35.3%)	28 (44.4%)
Several days a week	2 (13.3%)	4 (19.0%)	6 (17.6%)	8 (12.7%)
At least once a month	3 (20.0%)	4 (19.0%)	6 (17.6%)	10 (15.9%)
Less than once a year	1 (6.7%)	5 (23.8%)	6 (17.6%)	12 (19.0%)
Never	4 (26.7%)	2 (9.5%)	4 (11.8%)	5 (7.9%)
Experience with VR				
Everyday	0 (0.0%)	0 (0.0%)	1 (2.8%)	1 (1.8%)
Several days a week	0 (0.0%)	0 (0.0%)	3 (8.3%)	0 (0.0%)
At least once a month	6 (40.0%)	1 (5.0%)	6 (16.7%)	9 (16.1%)
Less than once a year	8 (53.3%)	11 (55.0%)	18 (50.0%)	29 (51.8%)
Never	1 (6.7%)	8 (40.0%)	8 (22.2%)	17 (30.4%)

Each group of participants was then required to complete an initial practice session to familiarize themselves with certain environments. The activity area included the indoor area shown in Fig. 4, with participants starting from the Starting Point in the waiting room and moving towards the Initial Point, in the meeting room. As shown in Fig. 4(a), upon reaching the Initial Point, each group was tasked with familiarizing themselves with the entire room's environment. After becoming fully acquainted with the surroundings, participants were instructed to return to the Starting Point via Exit A, a step designed to familiarize them with Exit A and its adjacent environment. Consistent with previous studies (Kinateder et al., 2018; Lovreglio et al., 2022), our experiment employed this method to ensure participants were familiar with Exit A before the evacuation scenario. As shown in Fig. 4(b), participants then experienced all eight fire evacuation scenarios listed in Table 2 in a random order. Each evacuation scenario began with an alarm trigger, signaling the onset of the evacuation. After the alarm sounded, NPCs in the environment began their evacuation, creating a dynamic evacuation atmosphere. Specific exits, were intermittently filled with smoke to simulate visibility reduction and increased urgency. During evacuation, pedestrians may split into subgroups (blue, evacuating toward the middle exit) and individuals (red, evacuating toward the bottom exit), as shown in Fig. 4(b). Evacuation succeeds and the scenario is completed when both subgroups and individuals reach the exit. After completing all eight scenarios, they filled out a post-experiment questionnaire, which included subjective feedback on their virtual experience. Following the questionnaire, participants took a 3-minute break and were observed to ensure they did not experience any discomfort before leaving (Lu et al., 2022).

2.4. Participants

A total of 140 participants were recruited through flyers and posters distributed in Tsinghua University to take part in the experiment, with most of them being university students and staff. All participants were randomly allocated to one of the four experimental configurations

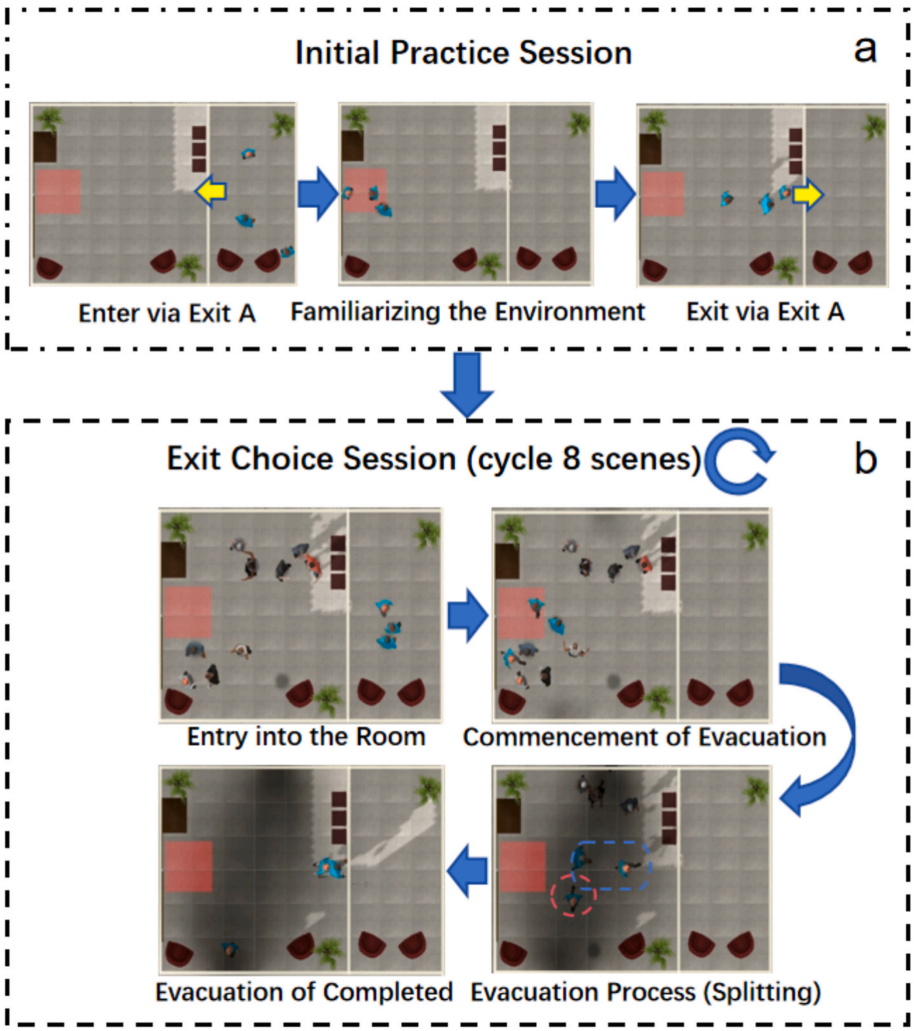

Fig. 4. Experimental procedure of evacuation including initial practice and exit choice sessions.

(group size 1, group size 2, group size 3 or group size 4). Due to network disconnections, the data points from ten participants were invalid. A posttest attention-check question filtered out another three data points. As a result, the final sample size consisted of 127 participants, 6 were staff members, and the rest were university students. They were organized into 51 groups, all of which were valid for data analysis. Specifically, the distribution was as follows: 15 participants were allocated to the group size 1 configuration, 20 to group size 2, 36 to group size 3, and 56 to group size 4. Table 3 outlines the demographics of the valid participants for each group size.

Since the assumption of a Chi-square test (at least 80% of the cells had an expected count larger than five) was violated, Fisher-Freeman-Halton exact tests were employed and confirmed that there was no significant difference among the three groups in terms of frequency of practice in a fire drill ($p = 0.24$), frequency of playing video games ($p = 0.80$), and experience with VR ($p = 0.07$).

The sample used in this work allows for the collection of 1016 observations. This number of observations is well above the minimum required when using the Hsieh method for power analysis. In fact, when assuming an alpha of 0.05, a power of 0.8, a sample proportion and ρ2 of 0.5, the required number of observations to detect a small effect size (OF = 1.2) is 306 (Chow et al., 2007; Hsieh et al., 1998). We chose to exceed this minimum sample size to ensure greater statistical reliability and robustness in our findings. By increasing the sample size, we could better capture potential variability within the data and reduce the margin of error, thus enhancing the validity of the conclusions drawn from our analysis.

2.5. *Measurements*

To ensure consistency in data collection and analysis, the core terms related to group dynamics (i.e., individuals post-splitting and subgroups) were applied as defined in Section 2.1.

A standardized, data-driven protocol was developed to objectively identify group splitting events, using behavioral data recorded by two complementary systems (Section 2.2): the VR system's automatic trajectory recorder (sampling rate: 1 Hz, capturing real-time coordinates of participants) and EV Capture (top-view screen recordings, enabling visual verification). Two mutually exclusive criteria were applied to cover all potential splitting scenarios:

Criterion 1: Exit choice divergence: For an initial group, if members selected different evacuation exits (labeled A, B, C in Fig. 1) during a trial, the group was deemed to have split.
Criterion 2: Trajectory separation: For members who did not exhibit exit choice divergence, splitting was determined by real-time Euclidean distance between original group members. If the distance between any two members reached ≥1.5 m and persisted until reaching the exit, the group was classified as split.

To minimize subjective bias, two independent researchers were trained to annotate splitting events using the EV Capture recordings and trajectory data. Discrepancies (e.g., ambiguous separation duration) were resolved through joint review of the raw recordings. Inter-rater reliability was quantified using Cohen's kappa coefficient ($\kappa = 0.89$, $p < 0.001$), indicating high consistency and validating the protocol's objectivity.

Several clusters of items were employed to collect information from the participants regarding: (1) Demographics. (2) Previous Experience (e.g., frequency of practicing in fire drills, frequency of playing video games, and experience with VR). (3) Level of familiarity. (4) Level of realism. (5) Level of urgency.(6) Validity. The pretest questionnaire encompasses clusters 1–3. The posttest questionnaire encompasses clusters 3–6 These clusters of items are described in more detail as follows.

The demographics cluster includes two items to collect participants' characteristics: gender and age. This information is gathered to ensure comparability among groups with different group sizes.

The previous experience cluster consists of four items that gather data on participants' past fire drill experiences and familiarity with video games and VR. This information is collected to ensure that different groups have comparable previous experiences, thereby minimizing any potential experience bias.

The familiarity cluster includes five items to assess the degree of familiarity among group members, drawing on methods from (Adams et al., 2005; Carlson & Zmud, 1999). The familiarity questions comprise the following four yes/no questions and one rating question, with the yes/no questions designed to reduce subjectivity and bias in the rating question: (1) Do you know the code names of all your companions? (2) Are you familiar with the virtual avatars of all your companions? (3) Do you know the hometowns of all your companions? (4) Are you familiar with the hobbies and interests of all your companions? and (5) Please rate your level of familiarity with your companions. Ratings are collected using a 7-point Likert scale (1 = "strongly disagree", 7 = "strongly agree"), and answers are averaged to form the final score for each participant in a group.

The realism cluster includes three items to assess the realism of the VR experience. It is measured using the items used by (Lovreglio et al., 2022). The realism questions include: (1) The virtual room was accurate/realistic. (2) The virtual smoke was accurate/realistic. (3) The virtual people were accurate /realistic.

The urgency cluster consists of one item used to assess participants' subjective perception of the urgency level during the VR emergency. It is measured using the items used by (Lovreglio et al., 2022). The urgency question is: (1) I felt the urgency to act/do something during the virtual emergency.

The validity cluster includes one item designed to assess the similarity between participants' behaviors in the VR emergency and their behaviors in real-life situations. It is measured using the items used by (Lovreglio et al., 2022). The validity question is: 1) I would act the same way in real life during the evacuation emergency.

2.6. Discrete choice modelling

The use of discrete choice models is well established in evacuation research to model exit choices due to its simplicity and effectiveness (Lovreglio, 2016). This model assumes that decision makers select the option that maximizes their perceived utility, which can be represented as a linear combination of relevant factors, such as exit distance and environmental conditions. Numerous studies have successfully applied MNL to study exit choices exit choice behavior during emergencies (Haghani & Sarvi, 2017; Kinateder et al., 2018; Gao et al., 2023; Zhu et al., 2023).

These models are grounded in the assumption that each decision-maker q assigns a utility $U_{q,i}$ to each available choice alternative i. The utility consists of a measurable component $V_{q,i}$ and a random component $\varepsilon_{q,i}$, as shown in Eq. (1) (Hensher et al., 2015; Ortúzar & Willumsen, 2011):

$$U_{q,i} = V_{q,i} + \varepsilon_{q,i} \tag{1}$$

This formulation accounts for two important aspects: first, that individuals with identical characteristics facing the same set of choices may select different alternatives; and second, that some individuals may not always choose what seems to be the most optimal option.

The measurable component is typically assumed to be a linear function of the factors $X_{q,i,j}$ perceived by the decision-maker q that affect the choice of alternative *i*. This is expressed in Eq. (2):

$$V_{q,i} = \sum_j \beta_{i,j} X_{q,i,j} \tag{2}$$

where $X_{q,i,j}$ are the known values of the factors j perceived by the decision-maker q affecting the choice for the alternative *i*. The $\beta_{i,j}$ are parameters weighting the preferences of the decision-makers related to the factors j. $\beta_{i,j}$ are the parameters to estimate.

In this study, the Multinomial Logit Model was employed, as it is a widely used discrete choice model known for its simplicity and practicality (Bourguignon et al., 2007; Darwish et al., 2024). To estimate the MNL model parameters, this study employed Maximum Likelihood Estimation (MLE), a standard and robust technique for discrete choice models. MLE identifies the set of $\beta_{i,j}$ values that maximize the "likelihood" of observing the actual exit choices made by participants in the experiment—i.e., the probability that the model, given a set of parameters, would reproduce the observed data. The foundation of MLE for MNL models lies in the closed-form solution for the probability $P_{q,i}$ that decision-maker q chooses exit i, which is derived under the assumption that the random component $\varepsilon_{q,i}$ follows an independent and identically distributed (i.i.d.) extreme value Type I distribution. This probability is:

$$P_{q,i} = \frac{exp(V_{q,i})}{\sum_k exp(V_{q,k})} \tag{3}$$

This equation can be employed to build a likelihood function, which is then utilized to estimate the $\beta_{i,j}$ parameters by identifying the parameter combination that maximises the likelihood function under the assumption of constant parameter (or random parameters in the case of a mixed logit model). MLE optimization was performed in R Studio using the mlogit package (Croissant, 2008), which employs the BFGS algorithm to iteratively adjust $\beta_{i,j}$ (starting from 0) until convergence (tolerance $= 10^{-6}$).

The dataset for MLE included 1016 valid observations (127 participants $\times$ 8 scenarios), with variables defined and captured objectively via the multi-user VR system to avoid subjective bias. The dependent variable (exit choice of decision-maker q) was automatically recorded by the VR system as the exit through which the individual or all members of a subgroup crossed the virtual boundary and completed evacuation—this operationalization defines the "exit decision moment" as the instant of evacuation completion. For independent variables: (1) NP (number of NPCs using the exit), DIST (distance to the exit), SMOKE (presence of smoke), and FAM (familiarity with the exit) were developed using a statistical Efficient Design as referenced in related work (Lovreglio et al., 2022). These predefined values were loaded automatically by the VR system to ensure consistency across trials; (2) GM (number of group members who chose the exit) was computed by the VR system in real time, counting how many original group members had already completed evacuation through the same exit before the current decision-maker.

It is important to address the Independence of Irrelevant Alternatives (IIA) assumption inherent in the multinomial logit model. As foundational literature explains, this property is a consequence of the assumption that the unobserved utility components (the error terms, $\varepsilon_{q,i}$) are independent and identically distributed (Train, 2009; Hensher et al., 2015). A common concern about IIA violation emerges when

Table 4
Estimated parameters for Eq. (4).

Parameter	Estimate	SD	Z-value	p-value
βnp	0.007	0.012	0.657	0.511
$\beta dist$	−0.341	0.059	−5.799	< 0.001
βsm	−1.549	0.132	−11.773	< 0.001
βfam	0.790	0.159	4.979	< 0.001
βgm	0.014	0.085	0.164	0.869
$\beta np1$	0.064	0.028	2.311	0.021
$\beta dist1$	0.519	0.149	3.480	0.001
$\beta sm1$	1.055	0.252	4.190	< 0.001
$\beta fam1$	−0.617	0.349	−1.768	0.077
βgm1	0.392	0.214	1.832	0.067
βnpG	0.026	0.020	1.291	0.197
$\beta distG$	−0.044	0.105	−0.423	0.673
βsmG	−1.326	0.247	−5.358	< 0.001
$\beta famG$	−0.183	0.278	−0.659	0.510
βgmG	0.805	0.238	3.389	< 0.001

unobserved factors (e.g., interdependencies between exit choices driven by social interactions) are embedded in the error terms, leading to spurious correlations between alternatives. However, this risk is significantly mitigated in our study through explicit modeling of key interdependent factors.

As outlined in Section 2.1 (Table 1) and reflected in the utility function (Eqs. (5)–(8), we incorporated two critical variables to capture social interaction effects: (1) GM (number of group members who have chosen the exit), which quantifies the influence of real peers' decisions; and (2) NP (number of NPCs using the exit), which captures crowd-induced herding behavior. These variables directly target the "interdependence between pedestrian choices and others' decisions"—the core source of potential IIA violation highlighted in theoretical discussions. By quantifying these effects in the systematic utility component ($V_{q,i}$) via estimated parameters (e.g., βgmG = 0.805, Table 4), we ensure that social interaction-driven interdependence is not hidden in the error terms.

Furthermore, environmental (DIST, SMOKE) and individual (FAM) factors are fully included to cover non-social sources of exit choice influence, leaving minimal unobserved variation in $\varepsilon_{q,i}$. This design ensures that the remaining error terms are unlikely to introduce significant interdependence across exit alternatives, supporting the validity of the MNL model in our study.

3. Results

Multifold analyses of the experimental results are conducted and reported in this section. Section 3.1 examines the group splitting patterns observed during the experiment. Section 3.2 presents the results of the exit choice model developed in this study, followed by Section 3.3, which provides a detailed sensitivity analysis of the key parameters influencing the exit decisions. Finally, Section 3.4 highlights the participants' feedback on the virtual experience, including their perceptions of realism and ecological validity.

$$\begin{gathered} U_i = (\beta np + C1 \times \beta np1 + C2 \times \beta npG)NP_i + (\beta dist + C1 \times \beta dist1 + C2 \times \beta distG)DIST_i \\ +(\beta sm + C1 \times \beta sm1 + C2 \times \beta smG)SMOKE_i + (\beta fam + C1 \times \beta fam1 + C2 \times \beta famG)FAM \\ +(\beta gm + C1 \times \beta tgm1 + C2 \times \beta gmG)GM_i \\ i = A, B, C \end{gathered} \tag{4}$$

3.1. Group splitting

Groups of 2–4 individuals typically maintain cohesion in normal situations. However, in emergency scenarios, weak social bonds may be insufficient to maintain cohesion, leading to group splitting (Wu & Zheng, 2024). In our study, we classify those separated after a group split as either individuals or subgroups. Group splitting pattern refers to the specific arrangement of individuals and subgroups that emerge after a group splits such as forming smaller units with mixed sizes. After a split, various group splitting patterns may emerge, including groups that remain intact. The tendencies of a group to split into smaller subgroups or remain as a single group were analyzed, as shown in Fig. 5.

Our experimental results revealed distinct patterns of splitting behavior across groups of different sizes. As shown in Fig. 5, these percentages represent the likelihood of each splitting pattern occurring within groups of a specific size across all evacuation trials (i.e., occurrences calculated based on total group number * trials). Specifically, two-person groups exhibited a strong tendency to split, with 71.25% (n = 57) of the groups splitting into two individuals during the evacuation. Only 28.75% (n = 23) of the two-person groups remained together. The three-person groups also showed a tendency to split into smaller units. The most frequent outcome was a split into a "1 two-person subgroup & 1 individual" configuration, occurring in 39.58% of the cases (n = 38). Only 28.75% (n = 23) of the three-person groups remained together. The four-person groups exhibited diverse splitting patterns. The most frequent outcome was a split into a "1 two-person subgroup & 2 individuals" configuration, occurring in 50.45% of the cases (n = 56), indicating that individuals often formed pairs. In 24.32% (n = 27) of the instances, the four-person groups were completely split into independent individuals, while only in 1.80% (n = 2) of the instances the groups remained intact throughout the evacuation. Other notable configurations included a "1 three-person subgroup & 1 individual" split (18.02%, n = 20) and a "2 two-person subgroup" split (5.41%, n = 6).

We tested whether participants' prior VR experience confounded our key behavioral dependent variable: group splitting. A chi-square test of independence was performed between VR usage frequency (Table 3) and the incidence of group splitting. The result was not statistically significant, $\chi2 = 1.676$, p = 0.195, indicating that prior VR proficiency did not have a significant association with group splitting behavior.

3.2. Exit choice model

In this study, we proposed a multinomial logit model formulation using the 1016 choice observations collected from the experiment. The model specification is illustrated in Eq. (4). We estimated the parameters $\beta_{i,j}$ weighting the impact of NP (NPCs using the exit), DIST (distance to exit), SMOKE (presence of smoke), FAM (familiarity with exit), and GM (groupmates' decisions). The parameters were all treated as generic across the three exits, meaning that we did not estimate alternative-specific parameters (Lovreglio et al., 2022).

To investigate the impact of participants' initial unawareness of the fire evacuation, we compared their behavior during the first trial to those during subsequent trials. We also sought to determine whether group or individual dynamics influenced their repeated decisions. To achieve these, we developed discrete choice models based on Eq. (4). Specifically, for individual decisions during the first evacuation, we set

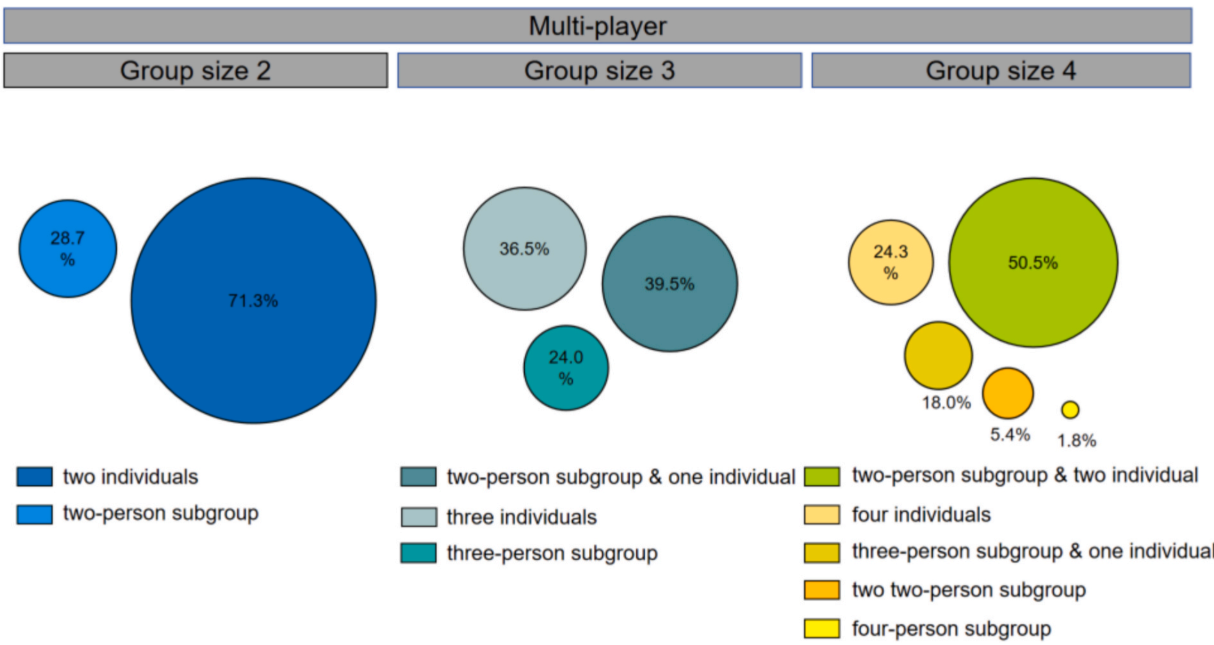

Fig. 5. The distribution of group splitting patterns for group sizes of 2–4.

C1 = 1 and C2 = 0. For subgroup decisions during the first evacuation, we set C1 = 1 and C2 = 1. For later evacuation trials, we set C1 = 0 and C2 = 0 for individual decisions, and C1 = 0 and C2 = 1 for subgroup decisions. Additionally, the parameters $\beta_{i,j}1$ represent the weights for individual decisions in the first trial, while $\beta_{i,j}G$ denote the weights for group decisions.

The estimated parameters for Eq. (4) are listed in Table 4. Only parameters with a p-value <0.05, indicating a significant effect on participants' decisions, were included in the calculations of the detailed formulas for the four scenarios—individual decision (first trial), subgroup decision (first trial), individual decision (subsequent trial), and subgroup decision (subsequent trial)—corresponding to Eqs. (5)–(8).

Individual decision (first trial):

$$U = 0.064np + 0.178dist - 0.494sm + 0.790fam \tag{5}$$

Subgroup decision (first trial):

$$U = 0.064np + 0.178dist - 1.82sm + 0.790fam + 0.805gm \tag{6}$$

Individual decision (subsequent trial):

$$U = -0.341dist - 1.549sm + 0.790fam \tag{7}$$

Subgroup decision (subsequent trial):

$$U = -0.341dist - 2.875sm + 0.790fam + 0.805gm \tag{8}$$

According to Eqs. (5)–(8), we observed that in individual decision-making, smoke, familiarity, and distance significantly influenced participants' exit decisions. In contrast, in subgroup decision-making, in addition to smoke, familiarity, and distance, decisions were also significantly influenced by the choices of other group members, exhibiting a tendency of the participants to follow others. Among all decision factors, smoke was the most significant, with a more pronounced impact than individual decision. Finally, comparing the first individual or subgroup decision with subsequent evacuation decisions, it is evident that during the first decision, both individual and subgroup decisions were significantly influenced by the choices of NPCs in the scenario, demonstrating a conformity effect. This conformity effect refers to participants aligning their decisions with others, likely due to uncertainty or the assumption that following others leads to safer outcomes, especially in high-pressure situations.

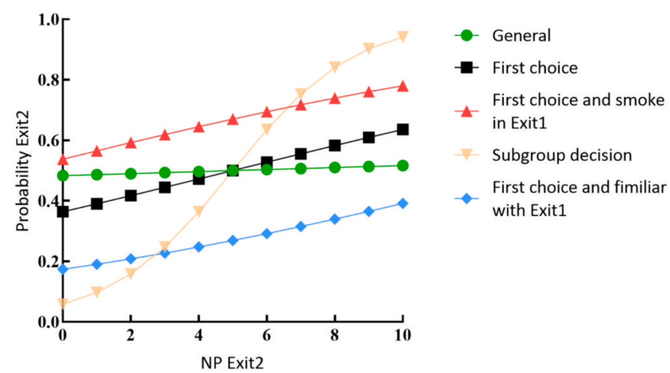

Fig. 6. Sensitive analysis of the proposed model.

3.3. Sensitivity analysis

Sensitivity testing was conducted to examine the influence of various factors on the probability of selecting an exit during evacuation, as shown in Fig. 6. This analysis built upon the discrete choice model outlined in Equation (1), focusing on key parameters such as the number of people at an exit, familiarity with the exit, group decision-making, and whether the decision was the initial choice. To simplify the analysis, we considered only two possible exits, assuming both exits were equidistant from the evacuees, with five individuals initially choosing Exit 1. Fig. 6 depicts the probability of selecting Exit 2 across various scenarios, with NP ranging from 0 to 10 at Exit 2. The conditions included a baseline condition without specific influences (green line), a first-choice condition (black line), a condition where smoke was present at Exit 1 during the first choice (red line), a group decision condition (orange line), and a condition where evacuees were familiar with Exit 1 (blue line).

The results showed a distinct trend: as the number of people at Exit 2 increases, so does the probability of evacuees choosing Exit 2 increased across all conditions. Compared to the baseline condition (green line), the first-choice condition (black line) showed a preference for more populated exits, highlighting a tendency to follow others during emergencies. The presence of smoke at Exit 1 during the first decision (red line) increased the probability of choosing Exit 2 by approximately 20% compared to the baseline condition (black line), emphasizing the significant role of environmental factors. Conversely, familiarity with Exit 1 (blue line) reduced the probability of selecting Exit 2 by approximately 23% compared to the condition where Exit 1 was unfamiliar (black line). The subgroup decision condition (orange line) exhibits a unique pattern, with a sharp increase in the likelihood of choosing Exit 2 as NP approaches 10, demonstrating the significant influence of social influence on exit selection during evacuation.

3.4. Subjective questionnaires

In our experiment, we assessed group members' familiarity with and participants' perceptions of the VR system's realism, urgency, and validity. Familiarity impacts group exit choice decision making, while realism and urgency measure the simulation's accuracy in reflecting real-life conditions. Validity ensures that participants' actions in the simulation mirror potential real-world behavior.

Familiarity was a controlled variable. A one-way ANOVA analysis was used to investigate the effects of group size on the level of familiarity within each group. The results, illustrated in Fig. 7, revealed no significant differences in familiarity across groups of sizes 2, 3 and 4.

After VR experiment, participants provided feedback on the system and the realism of the experience using a 7-point Likert scale. The results, shown in Fig. 8, gauge participants' perceptions of the realism of the virtual environment, smoke, and people (Realism1, Realism2, Realism3), the urgency felt during the virtual emergency (Urgency), and the validity of their actions as reflective of how they would respond in a real evacuation (Validity). For the realism of the virtual room, smoke, and people (Realism1, Realism2, Realism3), the median responses were above the midpoint of the scale, suggesting that participants found the VR simulation to be relatively realistic. It is interesting to note that the responses for the realism of the virtual environment (Realism1) had a slightly higher median value than the other two realism dimensions. Regarding the urgency, the median response was slightly lower but still above the scale's midpoint, indicating that participants felt a moderate sense of urgency to act during the VR emergency. For the validity of their actions, the participants' responses indicate that they believed their actions in the virtual scenario were similar to those they would take should they be caught in a real fire emergency. Overall, these results suggested that the VR simulation provided a comprehensive immersive experience that effectively elicited a sense of urgency.

4. Discussion

This study extends previous research on exit choice during evacuations by exploring decision-making among multiple participants within a VR environment, allowing group members a choice to evacuate together or individually. This flexible design better reflects real-world

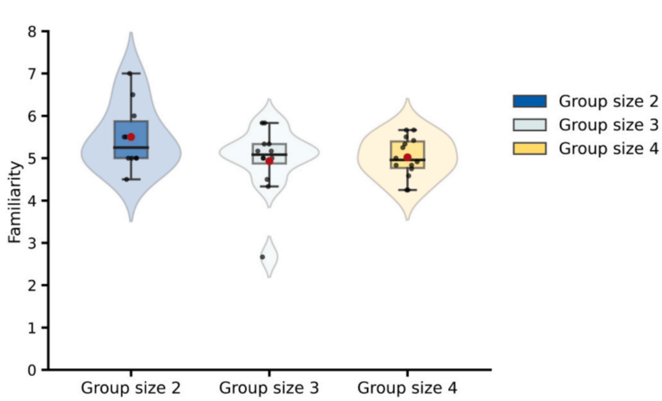

Fig. 7. Participants' scores on familiarity with group members.

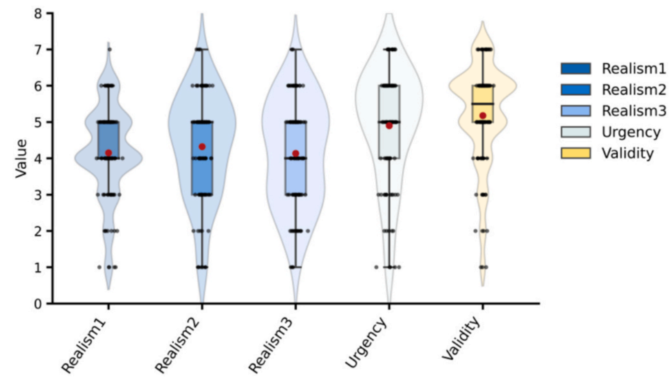

Fig. 8. Participants' scores on realism, urgency and validity.

dynamics, where social bonds influence, but do not entirely dictate, evacuation behavior. The findings offer critical insights into the decision-making processes of both individuals and subgroup, highlighting the complex interplay between social influences and individual judgement.

To our knowledge, this is the first study to implement multi-user VR experiments to explore exit choices during evacuation, introducing a new approach to studying crowd dynamics and behavior. Prior studies have relied on field experiments, which, while insightful, face significant ethical and safety limitations. Additionally, single-user VR studies often fail to capture the complex dynamics and social aspect of crowd behavior. By introducing a multi-user VR simulation paradigm, our study overcomes these limitations, offering a cost-efficient and ecologically valid platform for systematically studying group evacuation dynamics. As noted by Haghani (2023), while field studies offer high realism, controlled experiments (like VR) are essential for establishing causal relationships—such as the specific impact of smoke or social influence on subgroup decisions—which are often impossible to isolate in uncontrolled natural settings.

4.1. Group splitting characteristics

While previous studies have explored group decisions, they often assume groups remain intact throughout the evacuation process (Haghani et al., 2019). The possibility of group members splitting into temporary subgroups has largely been overlooked (Kerth, 2010). Our experiment addressed this gap by analyzing how groups with different sizes behave in evacuation scenarios, particularly focusing on the phenomenon of group splitting. The results revealed that group splitting was prevalent across all group sizes. It is important to note that the groups represented in this study were more akin to groups of friends or acquaintances, rather than tightly bonded family units. In evacuation scenarios involving friends or colleagues, the social ties are strong enough to influence behavior but not as rigid as in family groups, allowing for more flexibility in decision-making (Li et al., 2020; Sadri et al., 2021). Most notably, in three-person and four-person groups, two-person subgroups emerged as the dominant configuration. The observed high frequency of group splitting, especially into two-person subgroups, could be possibly due to the high costs of reaching a consensus under factors such time constraints, individual differences in information processing, or conflict among group members (Conradt & Roper, 2003, 2005, 2009; Franks et al., 2003; Haghani et al., 2019). Group splits occur when the costs of achieving consensus outweigh the benefits of collective action (Biro et al., 2006; Kerth, 2010). In larger groups, the increase in shared information enhances the groups' collective interests, but the complexity of decision-making also rises, demanding more coordination. This becomes even more pronounced during emergencies, when decisions must be made under pressure. As a result, splitting into smaller subgroups, particularly pairs, strikes a balance between maximizing group benefits and minimizing the costs of reaching a consensus. In three-person and four-person groups, the formation of two-person subgroups allows group members to retain some social benefits while reducing the difficulty of coordination.

However, in two-person groups, the advantages of shared information are more limited compared to larger groups, and reaching a consensus can be challenging. If the two members disagree, there is no majority to break the tie, which can lead to a stalemate. In such situations, the perceived costs of maintaining group cohesion may outweigh the benefits, prompting individuals to act independently rather than risking delay. This likely explains the high likelihood (71.25%) of splitting in two-person groups, as individuals seek to maximize their chances of survival when consensus is unattainable.

4.2. Comparative analysis of exit choices

After group splitting, several key factors influenced the exit choice of participants, both within subgroup and when acting independently. Our study examined four primary factors: familiarity, number of NPCs using the exit, smoke, and exit distance. These factors have been identified in prior research as comprehensive and critical for understanding the trade-offs participants make during an evacuation (Lovreglio et al., 2022).

Our analysis first examined the factors influencing individual decision-making following group splitting. Compared to previous studies (Duives & Mahmassani, 2012; Gao et al., 2022), our approach incorporates multiple factors influencing exit choice during fire evacuations. We found that individuals primarily considered smoke, familiarity, and exit distance, with smoke identified as the most significant factor. This finding aligns with previous research highlighting the strong influence of smoke on pedestrians' exit choices during evacuations (Gao et al., 2022; Lovreglio et al., 2022). Familiarity also significantly impacted exilt selection, consistent with Sime's affiliative model (Sime, 1983) and supported by prior empirical studies (Fridolf et al., 2013; Kinateder et al., 2018). Additionally, participants showed a clear preference for the nearest available exit, corroborating prior studies that emphasize minimizing travel distance and perceived risk in emergencies (Lovreglio et al., 2022). Notably, unlike earlier works (Feng et al., 2021a; Lovreglio et al., 2022), we observed that once individuals split from a group, they were no longer influenced by NPCs using the exit. In other words, those who became independent after leaving a group displayed less susceptibility to the "herding behavior" than individuals who were alone from the start, indicating a stronger sense of autonomy once they had separated.

Moreover, we contrast subgroup decision-making with individual choices, highlighting two key distinctions. Firstly, while both were influenced by the aforementioned factors, including familiarity, distance, and smoke, subgroup decisions were additionally shaped by the influence of other group members' choices. This behavior aligns with the Social Influence Theory (Asch, 1961) and Social Identity Theory (Tajfel & Turner, 2004), which suggest that participants are more likely to conform to social norms and align with the choices of others when in a cohesive group. According to Drury et al. (2009), this tendency is especially strong when maintaining group unity is prioritized, particularly in ambiguous or high-pressure situations. In contrast, when group members made individual decisions, they did not exhibit the same tendency to follow other group members. This divergence could be attributed to the urgency and life-threatening nature of a fire evacuation, where individuals may prioritize their safety over group unity. In such cases, individuals rely on personal judgment, particularly when group decisions could lead to overcrowding or when safer exits are available. This aligns with previous research, which suggests that in high-risk scenarios, personal safety often overrides the tendency to conform (Kinateder et al., 2014).

Furthermore, our findings diverge from those of Haghani et al. (2019) in their study on group exit choice, and this discrepancy originates from a core methodological distinction: enforced group cohesion versus natural splitting. Haghani et al. (2019) conducted a lab-in-the-field experiment that mandated all group members to evacuate collectively (enforced cohesion), with groups comprising individuals presumably with strong social ties (e.g., kinship ties). The intentional prevention of splitting in their design led to their conclusion that "the exit selection mechanisms of groups are largely similar to those of individuals."In contrast, our multi-user VR experiment simulates weak-tie groups (i.e., acquaintances/colleagues; consistent with Wu & Zheng, 2024) and eliminates the constraint of enforced cohesion—thereby enabling the emergence of natural splitting, an occurrence common in real-world emergencies (Cornwell & Ho, 2022; Feliciani et al., 2023).

This methodological difference has critical theoretical implications. By ignoring splitting, Haghani et al. (2019) could not observe the unique decision dynamics of post-split subgroups. Our study, however, finds that subgroups (unlike intact groups in Haghani et al.'s work) exhibit distinct patterns: (1) they are significantly influenced by peers' choices (GM variable, $\beta\_GM = 0.805$, Table 4), a social influence not observed in individual decision-making; (2) they are more sensitive to smoke (smoke coefficient: -2.875 for subgroups vs. -1.549 for individuals, Eqs. (7) and (8), a hazard awareness amplified by collective information processing. These findings extend Haghani et al.'s (2019) conclusion by showing that group exit-choice mechanisms are not universally similar to individuals—instead, they depend on whether groups remain intact or split, and on the strength of social ties (strong ties in Haghani et al. vs. weak ties in our study). This nuance is critical for evacuation theory, as it highlights the need to distinguish "intact group decision-making" from "split-group decision-making" when modeling crowd behavior.

Furthermore, our findings differ from Haghani's earlier study on group exit choice (Haghani et al., 2019), which concluded that the exit selection mechanisms of groups are largely similar to those of individuals. A likely explanation for this discrepancy is that Haghani's study enforced group evacuation, thereby overlooking the possibility of group splitting. Understanding this duality is crucial for emergency management, as it highlights the need to differentiate between group influence and individual autonomy in evacuation scenarios. By distinguishing this, emergency planners can design more effective interventions, such as targeted guidance systems and dynamic signage, that account for both subgroup decisions and individual responses.

Secondly, smoke emerges as the most critical factor in subgroup decisions, with a more pronounced effect than in individual settings. This heightened sensitivity to smoke in group contexts may stem from the collective processing of information, where the visible threat amplifies the group's focus on safety. Previous research has shown that in high-risk situations, group members often prioritize clear and immediate dangers, such as smoke, over other factors, due to shared discussions and heightened risk perception (Kinateder et al., 2014). This collective emphasis on the most apparent hazard likely drives the stronger influence of smoke on group decision-making, as the group dynamic reinforces the urgency to avoid the visible threat.

Based on the comparison between Eqs. (7) and (8) with 3 and 4, we also observed notable learning effect, similar to that reported in (Lovreglio et al., 2022), where participants' decisions evolved as they became more familiar with the experiment. In the initial scenario, participants primarily relied on social cues, such as the presence of others, to guide their exit choices. However, as they progressed through subsequent scenarios, their focus shifted more towards smoke as a critical hazard, indicating an adaptation and refinement of their decision-making strategies. This learning effect highlights the dynamic nature of human behavior in emergency situations.

To address potential learning bias and clarify ecological validity, pre-training with neutral tasks was implemented in our study As part of the study's "familiarization stage," participants completed pre-training in a smoke-free virtual room: they practiced VR operation (controller use, avatar movement) and explored the experimental environment (waiting room + meeting room) without emergency cues or exit-choice demands.

This measure separated “equipment adaptation” from “emergency decision-making,” ensuring subsequent evacuation choices reflected genuine responses to emergency stimuli, not adaptation to VR operation.

4.3. Participants perceptions

Our results were further supported by participants’ subjective feedback detailed in Section 3.4 and Fig. 8, which indicated a relatively high level of ecological validity for the VR simulation. Based on the Likert-scale ratings, participants perceived the environment and fire emergency as reasonably realistic and believed their actions reflected how they would behave in a real-life evacuation. While these perceptions were subjective, the quantitative survey data demonstrated the potential of VR as a robust tool for studying human behavior.

4.4. Limitations and future research

This study bears several limitations that should be acknowledged. Firstly, it is noteworthy that all participants in this experiment were students and university staff, which may limit the generalizability of the model. Certain behaviors observed might be specific to younger demographics, and further investigation is warranted in future research. Additionally, although we used subjective questionnaires to measure familiarity between group members and obtained relatively high familiarity scores, it is crucial to recognize that the familiarity simulated in this experiment may not fully capture the strong, inherent social ties typical of family members. The groups in this study were designed to represent friends or acquaintances (reflecting weak-to-moderate social ties). This type of social bond inherently shapes their exit choice behavior. In contrast, groups with stronger social tie strength (e.g., families) are likely to exhibit distinct evacuation behaviors (e.g., reduced splitting, more collective decision-making)—a pattern consistent with prior work (Haghani et al., 2019). Thus, exploring how different levels and types of social tie strength modulate group splitting and exit choices remains a key direction for future research.

Furthermore, regarding the modeling approach, we employed the Multinomial Logit (MNL) model without random parameters. This choice was made for two primary reasons. First, it ensures methodological consistency and direct comparability with foundational studies in this field, such as Haghani et al. (2019), which utilized a similar standard logit structure to provide robust and interpretable behavioral coefficients. Second, the MNL model’s closed-form equations are computationally efficient and straightforward to implement into agent-based evacuation simulation software, facilitating the practical application of our findings. While Mixed Logit models can account for unobserved individual heterogeneity, our primary focus was on capturing the average behavioral trade-offs between individuals and subgroups rather than assessing heterogeneity itself.

Moreover, while our study identified factors influencing individual and subgroup evacuation decisions, it did not quantitatively model and predict why individuals might choose to make decisions alone or participate in subgroup decision-making. Future research could integrate physiological data, such as heart rate and stress levels, to offer objective insights into how emotional and physical responses shape individual versus subgroup decision-making in evacuation scenarios. This data could be used to develop models that explain the underlying reasons for these decision-making patterns under stress. Another limitation is that we barely observed group reorganization (e.g., scattered people re-forming small groups during evacuation). This is because the virtual experimental area was small—only a meeting room with exits 3.0–6.0 m away. Participants reached exits in about 10 s (per VR trajectory data), leaving no time for scattered people to re-group. To fix this, future studies will use larger, more complex virtual areas (e.g., office buildings, metro stations) with longer evacuation paths ($\geq$20 m). This will let us better observe group reorganization and its patterns. Also, to make group splitting easier to understand, we added an evacuation video in the Appendix. The video shows typical evacuation trials (for 2-, 3-, 4-person groups) with simple labels—like colored avatars for group members and timestamps for when splitting starts—to clearly show how groups split.

5. Conclusion

In conclusion, this study employs multi-user VR and discrete choice models to address a critical underexplored gap in evacuation research: the exit choice dynamics of subgroups and individuals following group splitting. Unlike existing evacuation studies that either focus on individual behavior or enforce artificial group cohesion, this work delivers targeted contributions to theory, methodology, and practical emergency management.

Theoretically, this research advances evacuation decision-making theory by quantifying behavioral discrepancies between post-split subgroups and individuals—a phenomenon previously noted but never systematically analyzed. Our findings confirm that group splitting is prevalent among weak-tie groups (e.g., acquaintances, colleagues), with distinct drivers of exit choice: individuals primarily rely on exit distance, familiarity with exits, and smoke presence; subgroups, by contrast, are additionally shaped by social influence from group members and exhibit heightened sensitivity to smoke. This nuance extends prior frameworks that treat “group behavior” as a homogeneous category, highlighting the necessity of distinguishing “intact group” from “split-group” dynamics in evacuation theory.

Methodologically, this research advances the field on two fronts. First, it pioneers the use of multi-user VR as a novel experimental platform for evacuation exit choice research, moving beyond its more established application as a training tool. Our platform’s specific design, which simulates natural group splitting (rather than forced cohesion), captures real-time interactions between subgroups and individuals to generate ecologically valid behavioral data. Second, we extend the application of discrete choice models to compare the trade-off processes of exit selection between post-split entities, providing quantitative parameters critical for calibrating and refining evacuation simulation models.

Practically, these findings offer actionable guidance for emergency management and building safety design. Planners must account for the divergent decision patterns of post-split populations: individuals—characterized by higher autonomy—require targeted guidance to mitigate risks of uncoordinated evacuation; subgroups—sensitive to both social cues and hazard signals—benefit from safety signage that balances social information (e.g., peer behavior) and clear hazard warnings. Furthermore, evacuation simulation models should integrate the phenomenon of group splitting and the distinct decision patterns of subgroups and individuals, rather than assuming groups remain intact, to improve the accuracy of safety risk predictions.

CRediT authorship contribution statement

Song Lu: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jieyu Chen:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Nan Li:** Writing – review & editing, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Zhenan Feng:** Writing – review & editing, Methodology. **Milad Haghani:** Writing – review & editing, Methodology. **Ruggiero Lovreglio:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank all participants for taking part in this study. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant No. 72374120. The authors from Tsinghua University are grateful for the support of the NSFC. Any opinions, findings, conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agency.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ssci.2026.107111.

Data availability

Data will be made available on request.

References

Adams, S.J., Roch, S.G., Ayman, R., 2005. Communication medium and member familiarity: the effects on decision time, accuracy, and satisfaction. Small Group Res. 36 (3), 321–353. https://doi.org/10.1177/1046496405275232.

Asch, S.E., 1961. Effects of group pressure upon the modification and distortion of judgments. In: Henle, M. (Ed.), *Documents of Gestalt Psychology*. University of California Press, pp. 222–236. https://doi.org/10.1525/9780520313514-017.

Benthorn, L., Frantzich, H., 1999. Fire alarm in a public building: how do people evaluate information and choose an evacuation exit? Fire Mater. 23 (6), 311–315. https://doi.org/10.1002/(SICI)1099-1018(199911/12)23:6%253C311::AID-FAM704%253E3.0.CO;2-J.

Biro, D., Sumpter, D.J.T., Meade, J., Guilford, T., 2006. From compromise to leadership in pigeon homing. Curr. Biol. 16 (21), 2123–2128. https://doi.org/10.1016/j.cub.2006.08.087.

Bode, N.W.F., Codling, E.A., 2013. Human exit route choice in virtual crowd evacuations. Anim. Behav. 86 (2), 347–358. https://doi.org/10.1016/j.anbehav.2013.05.025.

Bode, N.W.F., Holl, S., Mehner, W., Seyfried, A., 2015. Disentangling the impact of social groups on response times and movement dynamics in evacuations. PLOS ONE 10 (3). https://doi.org/10.1371/journal.pone.0121227.

Bourguignon, F., Fournier, M., Gurgand, M., 2007. Selection bias corrections based on the multinomial logit model: Monte Carlo comparisons. J. Econ. Surv. 21 (1), 174–205. https://doi.org/10.1111/j.1467-6419.2007.00503.x.

Carlson, J.R., Zmud, R.W., 1999. Channel expansion theory and the experiential nature of media richness perceptions. Acad. Manag. J. 42 (2), 153–170. https://doi.org/10.2307/257090.

Chen, J., Shi, T., Li, N., 2021. Pedestrian evacuation simulation in indoor emergency situations: approaches, models and tools. Saf. Sci. 142, 105378. https://doi.org/10.1016/j.ssci.2021.105378.

Chow, S.-C., Shao, J., Wang, H., 2007. Sample Size Calculations in Clinical Research, 2nd ed. Chapman & Hall/CRC, Boca Raton, FL.

Conradt, L., Roper, T.J., 2003. Group decision-making in animals. Nature 421 (6919), 155–158. https://doi.org/10.1038/nature01294.

Conradt, L., Roper, T.J., 2005. Consensus decision making in animals. Trends Ecol. Evol. 20 (8), 449–456. https://doi.org/10.1016/j.tree.2005.05.008.

Conradt, L., Roper, T.J., 2009. Conflicts of interest and the evolution of decision sharing. Philos. Trans. R. Soc., B 364 (1518), 807–819. https://doi.org/10.1098/rstb.2008.0257.

Cornwell, B., Ho, J.-M., 2022. Network structure in small groups and survival in disasters. Soc. Forces 100 (3), 1357–1384. https://doi.org/10.1093/sf/soab036.

Darwish, A.M., Almansour, M., Salah, A., Zagow, M., Saeed, K., Elkafoury, A., 2024. Sensitivity evaluation of machine learning-based calibrated transportation mode choice models: a case study of Alexandria City, Egypt. Transport. Res. Interdiscipl. Perspect. 24, 101052. https://doi.org/10.1016/j.trip.2024.101052.

Drury, J., Cocking, C., Reicher, S., 2009. Everyone for themselves? A comparative study of crowd solidarity among emergency survivors. Br. J. Soc. Psychol. 48 (Pt 3), 487–506. https://doi.org/10.1348/014466608X357893.

Duives, D.C., Mahmassani, H.S., 2012. Exit choice decisions during pedestrian evacuations of buildings. Transport. Res. Rec.: J. Transport. Res. Board 2316 (1), 84–94. https://doi.org/10.3141/2316-10.

Feliciani, C., Jia, X., Murakami, H., Ohtsuka, K., Vizzari, G., Nishinari, K., 2023. Social groups in pedestrian crowds as physical and cognitive entities: extent of modeling and motion prediction. Transport. Res. Part A: Policy Practice 176, 103820. https://doi.org/10.1016/j.tra.2023.103820.

Feng, Y., Duives, D.C., Hoogendoorn, S.P., 2021a. Using virtual reality to study pedestrian exit choice behaviour during evacuations. Saf. Sci. 137, 105158. https://doi.org/10.1016/j.ssci.2021.105158.

Feng, Y., Duives, D., Daamen, W., Hoogendoorn, S., 2021b. Data collection methods for studying pedestrian behaviour: a systematic review. Build. Environ. 187, 107329. https://doi.org/10.1016/j.buildenv.2020.107329.

Fisher, B.A., 1970. Decision emergence: phases in group decision-making. Speech Monogr. 37 (1), 53–66. https://doi.org/10.1080/03637757009375649.

Franks, N.R., Dornhaus, A., Fitzsimmons, J.P., Stevens, M., 2003. Speed versus accuracy in collective decision making. Proc. R. Soc. Lond. B 270 (1532), 2457–2463. https://doi.org/10.1098/rspb.2003.2527.

Fridolf, K., Nilsson, D., Frantzich, H., 2013. Fire evacuation in underground transportation systems: a review of accidents and empirical research. Fire Technol. 49 (2), 451–475. https://doi.org/10.1007/s10694-011-0217-x.

Gao, D., Lee, E.W.M., Lee, Y.Y., 2023. The influence of context effects on exit choice behavior during building evacuation combining virtual reality and discrete choice modeling. Adv. Eng. Inform. 57, 102072. https://doi.org/10.1016/j.aei.2023.102072.

Gao, D.L., Xie, W., Ming Lee, E.W., 2022. Individual-level exit choice behaviour under uncertain risk. Physica A 604, 127873. https://doi.org/10.1016/j.physa.2022.127873.

Grosshandler, W., Bryner, N., Madrzykowski, D., Kuntz, K., 2005. Report of the Technical Investigation of The Station Nightclub Fire.

Gwynne, S. M. V., Rosenbaum, E. R., 2016. Employing the hydraulic model in assessing emergency movement. In M. J. Hurley, D. Gottuk, J. R. Hall, K. Harada, E. Kuligowski, M. Puchovsky, J. Torero, J. M. Watts, & C. Wieczorek (Eds.), SFPE Handbook of Fire Protection Engineering, Springer, pp. 2115–2151. doi: 10.1007/978-1-4939-2565-0_59.

Haghani, M., 2023. The notion of validity in experimental crowd dynamics. Int. J. Disaster Risk Reduct. 93, 103750. https://doi.org/10.1016/j.ijdrr.2023.103750.

Haghani, M., Sarvi, M., 2017. Stated and revealed exit choices of pedestrian crowd evacuees. Transport. Res. Part B: Methodol. 95, 238–259. https://doi.org/10.1016/j.trb.2016.10.019.

Haghani, M., Sarvi, M., 2018. Crowd behaviour and motion: empirical methods. Transp. Res. B Methodol. 107, 253–294. https://doi.org/10.1016/j.trb.2017.06.017.

Haghani, M., Sarvi, M., Shahhoseini, Z., Boltes, M., 2019. Dynamics of social groups' decision-making in evacuations. Transp. Res. Part C Emerg. Technol. 104, 135–157. https://doi.org/10.1016/j.trc.2019.04.029.

Haghani, M., Sarvi, M., Shahhoseini, Z., 2020. Evacuation behaviour of crowds under high and low levels of urgency: experiments of reaction time, exit choice and exit-choice adaptation. Saf. Sci. 126, 104679. https://doi.org/10.1016/j.ssci.2020.104679.

Hensher, D.A., Rose, J.M., Greene, W.H., 2015. *Applied Choice Analysis*, second ed. Cambridge University Press. 10.1017/CBO9781316136232.

Hsieh, F.Y., Bloch, D.A., Larsen, M.D., 1998. A simple method of sample size calculation for linear and logistic regression. Statistics in Medicine 17 (14), 1623–1634.

Institute of Transport and Logistics Studies, 2007. Ngene—A Software Capability to Design and Generate Choice Experiments. University of Sydney.

Kerth, G., 2010. Group decision-making in fission–fusion societies. Behav. Process. 84 (3), 662–663. https://doi.org/10.1016/j.beproc.2010.02.023.

Kinateder, M., Warren, W.H., 2021. Exit choice during evacuation is influenced by both the size and proportion of the egressing crowd. Physica A 569, 125746. https://doi.org/10.1016/j.physa.2021.125746.

Kinateder, M., Müller, M., Jost, M., Mühlberger, A., Pauli, P., 2014. Social influence in a virtual tunnel fire – influence of conflicting information on evacuation behavior. Appl. Ergon. 45 (6), 1649–1659. https://doi.org/10.1016/j.apergo.2014.05.014.

Kinateder, M., Comunale, B., Warren, W.H., 2018. Exit choice in an emergency evacuation scenario is influenced by exit familiarity and neighbor behavior. Saf. Sci. 106, 170–175. https://doi.org/10.1016/j.ssci.2018.03.015.

Kuligowski, E.D., 2016. Human behavior in fire. In: Hurley, M.J., Gottuk, D., Hall, J.R., Harada, K., Kuligowski, E., Puchovsky, M., Torero, J., Watts, J.M., Wieczorek, C. (Eds.), *SFPE Handbook of Fire Protection Engineering*. Springer, New York, pp. 2070–2114. https://doi.org/10.1007/978-1-4939-2565-0_58.

Li, L., Ding, N., Ma, Y., Zhang, H., Jin, H., 2020. Social relation network and group behavior based on evacuation experiments. Physica A 545, 123518. https://doi.org/10.1016/j.physa.2019.123518.

Lovreglio, R., 2016. *Modelling decision-making in fire evacuation based on random utility theory*. Unpublished. https://doi.org/10.13140/RG.2.1.1695.5281/1.

Lovreglio, R., Borri, D., dell'Olio, L., Ibeas, A., 2014. A discrete choice model based on random utilities for exit choice in emergency evacuations. Saf. Sci. 62, 418–426. https://doi.org/10.1016/j.ssci.2013.10.004.

Lovreglio, R., Dillies, E., Kuligowski, E., Rahouti, A., Haghani, M., 2022. Exit choice in built environment evacuation combining immersive virtual reality and discrete choice modelling. Autom. Constr. 141, 104452. https://doi.org/10.1016/j.autcon.2022.104452.

Lu, L., Chan, C.-Y., Wang, J., Wang, W., 2017. A study of pedestrian group behaviors in crowd evacuation based on an extended floor field cellular automaton model. Transport. Res. Part C: Emerg. Technol. 81, 317–329. https://doi.org/10.1016/j.trc.2016.08.018.

Lu, S., Wang, F., Li, X., Shen, Q., 2022. Development and validation of a confined space rescue training prototype based on an immersive virtual reality serious game. Adv. Eng. Inf. 51, 101520. https://doi.org/10.1016/j.aei.2021.101520.

McGrath, J.E., 1991. Time, interaction, and performance (TIP): a theory of groups. Small Group Res. 22 (2), 147–174. https://doi.org/10.1177/1046496491222001.

Moreland, R.L., Levine, J.M., 2002. Socialization and trust in work groups. Group Process. Intergr. Relat. 5 (3), 185–201. https://doi.org/10.1177/1368430202005003002.

Ortúzar, J.D.D., Willumsen, L.G., 2011. *Modelling Transport*, 1st ed. Wiley. 10.1002/9781119993308.

Photon Pun2 Engine, 2018. https://doc.photonengine.com/pun/current/getting-started/pun-intro.

Rahouti, A., Lovreglio, R., Datoussaïd, S., Descamps, T., 2021. Prototyping and validating a non-immersive virtual reality serious game for healthcare fire safety training. Fire Technol. 57 (6), 3041–3078. https://doi.org/10.1007/s10694-021-01098-x.
Ren-Yong, G., Hai-Jun, H., 2010. Logit-based exit choice model of evacuation in rooms with internal obstacles and multiple exits. Chin. Phys. B 19 (3), 030501. https://doi.org/10.1088/1674-1056/19/3/030501.
Sadri, A.M., Ukkusuri, S.V., Ahmed, M.A., 2021. Review of social influence in crisis communications and evacuation decision-making. Transp. Res. Interdiscip. Perspect. 9, 100325. https://doi.org/10.1016/j.trip.2021.100325.
SÁndor, Z., Wedel, M., 2001. Designing conjoint choice experiments using managers' prior beliefs. J. Mark. Res. 38 (4), 430–444. https://doi.org/10.1509/jmkr.38.4.430.18904.
Sime, J.D., 1983. Affiliative behaviour during escape to building exits. J. Environ. Psychol. 3 (1), 21–41. https://doi.org/10.1016/S0272-4944(83)80019-X.
Tajfel, H., Turner, J.C., 2004. The social identity theory of intergroup behavior. In: Jost, J.T., Sidanius, J. (Eds.), *Political Psychology*, 0 ed. Psychology Press, pp. 276–293. https://doi.org/10.4324/9780203505984-16.
Train, K.E., 2009. Discrete Choice Methods with Simulation (Second Edition). Cambridge University Press.
Unity, 2022. Unity. https://unity.com/products/unity-engine.
Wu, W., Zheng, X., 2024. A systematic analysis of subgroup research in pedestrian and evacuation dynamics. IEEE Trans. Intell. Transp. Syst. 25 (2), 1225–1246. https://doi.org/10.1109/TITS.2023.3318417.
Xia, X., Chen, J., Zhang, J., Li, N., 2024. How the strength of social relationship affects pedestrian evacuation behavior: a multi-participant fire evacuation experiment in a virtual metro station. Transport. Res. Part C: Emerg. Technol. 167, 104805. https://doi.org/10.1016/j.trc.2024.104805.
Zhang, Y., Xie, W., Chen, S., Li, T., 2018. Experimental study on descent speed on stairs of individuals and small groups under different visibility conditions. Fire Technol. 54 (3), 781–796. https://doi.org/10.1007/s10694-018-0710-6.
Zhu, R., Becerik-Gerber, B., Lin, J., Li, N., 2023. Behavioral, data-driven, agent-based evacuation simulation for building safety design using machine learning and discrete choice models. Adv. Eng. Inf. 55, 101827. https://doi.org/10.1016/j.aei.2022.101827.