Virtual Reality Applications for Civilian Emergency Preparedness Training: A Bibliometric Review

Rashika Tasnim Keya
Western Norway University of Applied Sciences
Rashika.Tasnim.Keya@hvl.no

Ilona Heldal
Western Norway University of Applied Sciences
Ilona.Heldal@hvl.no

Daniel Patel
Western Norway University of Applied Sciences
Daniel.Patel@hvl.no

Pietro Murano
Oslo Metropolitan University
piemur@oslomet.no

ABSTRACT

This bibliometric review examines serious games (SG) supported by virtual reality (VR) technologies for civilian emergency preparedness training. It maps publication trends, research themes, and evaluation practices in 66 studies published from 2010–2024 using a PRISMA-ScR approach. The quantitative analysis was combined with qualitative coding to assess how the papers defined evaluation methodologies and assessed their effectiveness for VRSG. The results point to notably increased research activity after 2018. Fire and evacuation scenarios dominate, while earthquakes, floods/tsunamis/typhoons, and hazards such as landslides or extreme weather appear far less often. Most studies rely on in-game behaviour tracking and self-report questionnaires, with fewer controlled comparisons and very rare long-term retention assessments. The field is expanding and technically varied, but remains fragmented in outcome measures, reporting standards, and evidence linking in-game performance to real-world preparedness. Common methodological limitations include small samples, short-term evaluations, and limited validation of learning transfer.

Keywords

Virtual reality, Serious games, Emergency preparedness, Evaluation methods, Bibliometric analysis.

INTRODUCTION

Immersive technologies such as Virtual Reality (VR) and VR-based Serious Games (VRSG) are gaining increasing attention as potential tools for emergency preparedness training (EPT) (Abbas et al., 2023; Alshowair et al., 2024). Their interactive and immersive qualities allow learners to engage with realistic experiences, practice decision-making, and receive immediate feedback in ways that traditional training formats cannot easily provide (Kanade & Duffy, 2022; Tang et al., 2020). By simulating hazardous environments and high-stress situations, VR enables participants to practice decision-making, safety protocols, and hazard recognition in a safe, non-intrusive, and controlled setting (Chittaro & Buttussi, 2015; del Carmen Cardós-Alonso et al., 2024). Unlike traditional training methods, such as classroom instruction, tabletop exercises, or physical drills, VR-based environments can offer repeatable, scenario-driven experiences with adjustable difficulty, immediate feedback, and the ability to safely reproduce rare or dangerous conditions (Cha et al., 2012) exactly as many times as it is needed, or eventually by adjusting the level of difficulty to more challenging learning goals (Wijkmark, 2024). The benefit of serious game is increased engagement and allowing different population segments to train systematically (Susi, Johannesson, & Backlund, 2007). In this paper, VRSG refer to immersive VR applications that integrate game-based elements, such as interactive scenarios, goals, and feedback, for learning through emergency preparedness training rather than entertainment.

Historically, VR-based EPT applications, prototypes, and trials have concentrated on professional responders, notably firefighters and Emergency Medical Services (EMS) personnel, or Ambulance personnel (Heldal,

Wijkmark, & Pareto, 2016; Lochmannová et al., 2022). Research on VR for educating civilians, such as school students, workplace employees, and community members, has only recently started to grow (Feng et al., 2021). This shift reflects increasing recognition that members of the public are often the first individuals affected during emergencies and that their ability to recognize hazards, make decisions, and respond appropriately can influence safety outcomes before professional responders arrive (Paton, 2013).

Despite the expanding literature, important knowledge gaps remain between the possibilities for training civilians and the actual ability to implement new solutions to train them (Keya et al., 2025). Disaster scenarios vary widely (e.g., indoor fires, floods, earthquakes, and multi-hazard preparedness contexts), training objectives differ across studies, and the evaluation methods used to assess training outcomes are highly heterogeneous, making it difficult to define a common baseline. Demonstrating learning transfer in real disaster contexts is inherently difficult, as real emergencies are rare and unpredictable and cannot be experimentally staged with full ecological validity or genuine risk exposure. This makes it difficult to assess whether findings from different exercises and training contexts are comparable (Feng et al., 2018). Moreover, the balance between self-reported outcomes (e.g., perceived learning, confidence gains) and in-game objective behavioral measures (e.g., evacuation time, hazard mitigation actions) is inconsistent, raising questions about the extent to which current evaluation approaches capture skill transfer to real-world situations (Shaw et al., 2019). Preparedness in the public domain depends largely on voluntary engagement rather than mandated institutional drills, which differentiates civilian training from professional responder training.

Bibliometric analysis offers a systematic approach to map research developments, identify thematic trends, and highlight underexplored areas in the field (Donthu et al., 2021). While previous reviews have examined VR in emergency management more broadly (Alshowair et al., 2024; Lovreglio, 2020; Stefan, Mortimer, & Horan, 2023), only a few have combined bibliometric analysis with a focused investigation into evaluation practices, particularly for non-professionals in emergency. A bibliometric perspective allows examination not only of growth over time but also of how disaster types, research themes, and evaluation practices are distributed across the literature. Such mapping can clarify the maturity of the domain and reveal structural concentrations or gaps that may not be apparent through narrative review alone.

This study, therefore, conducts a bibliometric review of VR-based serious games and immersive VR applications for emergency preparedness training targeting civilians, covering publications from 2010 to 2024. The analysis maps publication patterns, thematic trends, disaster types addressed, and methodological approaches used within this body of research. Evaluation practices are examined as one analytical dimension within this broader bibliometric mapping, rather than as the sole focus of the study. By integrating quantitative bibliometric mapping with structured content analysis, this review aims to clarify how the field has developed over the past decade and to identify structural gaps that may guide future research on civilian emergency preparedness using VR technologies.

METHODS

Search Strategy and Data Sources

This bibliometric study aims to identify and analyse peer-reviewed publications on VR-based serious games for EPT targeting civilians, while mapping structural patterns, publication trends, and evaluation practices in this emerging interdisciplinary domain. Data were collected from **Scopus**, **Web of Science** (WoS), and **Google Scholar**. These repositories were selected to provide broad coverage of computer science, engineering, and applied social science research. For Scopus and Web of Science, the following search string was applied to titles, abstracts, and keywords: ("virtual reality" OR "VR" OR "HMD" OR "Head Mounted Display" OR "Virtual Environment") AND (evacuat* OR emergenc* OR disaster* OR incident*) AND ("serious games" OR "video games") AND (train* OR learn*) AND evaluat*. For **Google Scholar**, the search terms were adapted to account for platform limitations, using combinations of core keywords from the above string (e.g., "virtual reality serious game", "VR emergency preparedness training." Since Google Scholar does not support complex Boolean queries in the same way as structured databases, searches were conducted using combinations of the core keywords derived from the main search string. Results were filtered by publication year and relevance ranking, and the first set of results was manually screened to identify potentially relevant studies. Records retrieved through Google Scholar were treated as supplementary sources and were subjected to the same inclusion and exclusion criteria as records obtained from Scopus and Web of Science.

While immersive VR environments enabled by head-mounted displays (HMDs), as well as systems such as CAVEs, were explored as early as the 1960s and 1990s, these technologies were primarily limited to research laboratories and professional or military contexts (Buxton & Fitzmaurice, 1998). This paper focuses on publicly available consumer VR technologies.

The widespread consumer adoption of immersive VR did not occur until the mid-2010s, when Head-Mounted Displays (HMDs), especially the Oculus Rift, appeared in 2012, and VR became more of a consumer tool rather than a technology in research laboratories. This transition was driven by the commercial availability of affordable consumer-grade headsets, which significantly lowered technical and financial barriers for both users and developers (LaValle et al., 2014; Slater & Sanchez-Vives, 2016). To avoid missing earlier research on this technology and capture the period of increasing accessibility and adoption of immersive VR, the bibliometric analysis focuses on publications from 2010 to 2024. Additional factors influencing this timeframe include improvements in graphics pipelines, tracking technologies (e.g., motion and eye tracking), and the increasing frequency and visibility of disaster events such as fires, floods, and extreme weather (Székely et al., 2025).

To define how other researchers assess VRSG applications and take us closer to defining a good evaluation method for civilians, the four authors behind this paper agreed upon the following criteria:

Inclusion Criteria: We included research papers that met all the following conditions:

- **Topic Scope:** They present a VR-based serious game application for emergency preparedness training aimed at non-professional audiences (e.g., school children, community members, employees at different workspaces who need to train).
- **Empirical Evaluation:** Report empirical data from a user testing study or based on evaluating a VRSG training.
- **Publication Type:** Appeared as a peer-reviewed journal article, conference paper, or scholarly book chapter (published in English).

Exclusion Criteria: Studies were excluded if any of the following applied:

- They are focused solely on professional responder training (e.g., firefighters, paramedics) or on purely clinical/health-related emergency scenarios (outside the disaster preparedness context).
- Lacked a VR-based component (for instance, studies of emergency training that did not incorporate any VR/immersive environment).
- Where review papers or other secondary research (i.e., do not present original experimental or observational results).

While there are several literature reviews in the area, e.g., about learning by using immersive VR, in general or for training (Kaplan et al., 2020; Radianti et al., 2020), for emergencies regarding certain types of evaluations or specific issues, such as safety or understanding the influence of VR characteristics (Alshowair et al., 2024; Scorgie et al., 2024; Stefan, Mortimer, & Horan, 2023; Strojny & Dużmańska-Misiarczyk, 2023), or often for professionals (Williams-Bell et al., 2015; Zhu & Li, 2021) these reviews don't consider VRSG for civilians according to our knowledge today. These are used as inspiration for our review.

Screening and Selection Process

Search results were imported into Zotero, and duplicates were removed. Titles and abstracts were screened against predefined inclusion and exclusion criteria. During the initial screening phase, a subset of studies presented ambiguities regarding inclusion and exclusion criteria (e.g., unclear target populations, health-related disasters, and evaluation methods). These records were rechecked by revisiting their abstracts and, when necessary, full texts until a final inclusion decision was reached.

The literature identification and screening process was reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines (Tricco et al., 2018). Figure 1 presents the PRISMA flow diagram of identification, screening, and inclusion. In total, the process yielded a final dataset of N = 66 studies included in this review (after all exclusions were applied). This reduction from the initial 169 records reflects the sequential application of the inclusion and exclusion criteria summarized in Figure 1.

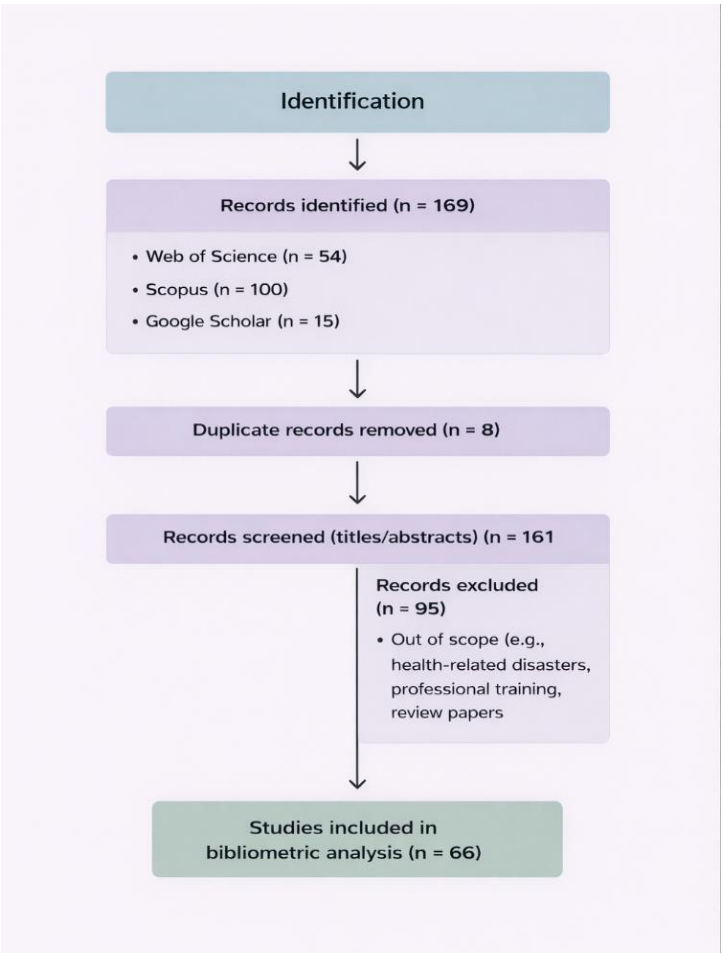

Figure 1. Summarizing Adapted PRISMA for Bibliometric Review flow chart for the Review Process.

Data Extraction, Coding, and Bibliometric Analysis

For each included study, bibliographic data were extracted, including publication year, source title, author affiliations, country of origin, and publication type. Additional study characteristics, such as the targeted emergency scenario (e.g., fire, earthquake, flood), evaluation method, and VR technology type, were also recorded. This process followed established bibliometric approaches for mapping research trends in emerging technology domains (Donthu et al., 2021).

In addition to bibliometric variables, each study was coded for its evaluation methods. Evaluation approaches were identified from abstracts and methods sections and inductively grouped into categories such as self-report measures, behavioural performance, knowledge tests, controlled comparisons, qualitative methods, physiological measures, and long-term follow-up. Because several studies applied more than one evaluation method, these categories were treated as analytical dimensions rather than mutually exclusive groups. A single study could therefore be assigned to multiple evaluation categories.

For the bibliometric analysis, the VOSviewer software (van Eck & Waltman, 2010) was being used to conduct quantitative mapping of publication trends, prominent sources, and keyword co-occurrence networks in the dataset. Where appropriate, additional analyses were conducted using the Bibliometrix package in R to complement the visualizations and statistical summaries of the publication data. We generated descriptive statistics to summarize core characteristics of the research (e.g., publication counts per year and distributions by venue or region) and created a keyword co-occurrence map to identify major thematic clusters and emerging topic areas within this field.

Finally, we qualitatively reviewed and synthesized observations from the included studies to highlight recurring limitations, trends in evaluation practices, and underexplored topics. These narrative insights (e.g., common challenges noted, gaps in study design or populations addressed) complement the quantitative bibliometric findings by providing context on how the field is evolving and where further research is needed.

RESULTS BASED ON THE BIBLIOMETRIC REVIEW

The bibliometric analysis revealed how research on VR-based emergency preparedness training has developed over time and across publication venues. This section summarizes the key trends identified across the 66 included studies, focusing on (1) how publication output has changed over time, (2) where this research is being disseminated, (3) what thematic areas dominate through keyword clusters, (4) which disaster types are most represented, and (5) how training effectiveness is evaluated. Together, these findings provide an evidence base for understanding the maturity, focus, and methodological balance of current research in this domain.

Publication Trends Over Time

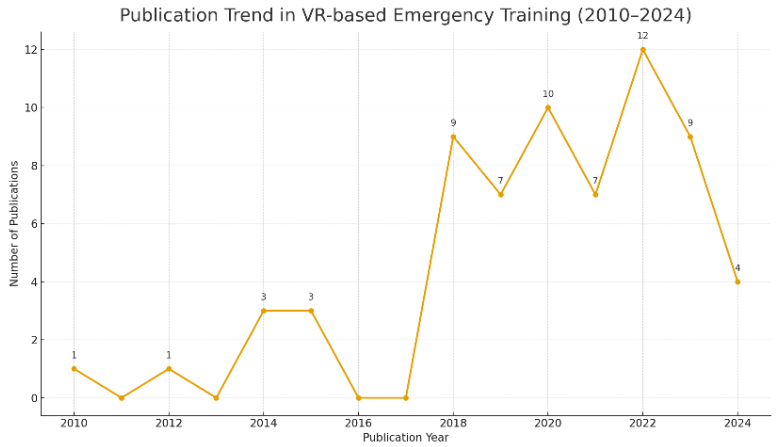

Figure 2. Publication Trends in VR-Based Emergency Training (2010–2024). Counts Reflect the Number of Indexed Items in our Dataset; Values for the Most Recent Years may be Incomplete due to Indexing Lag.

Figure 2 illustrates the annual publication trends on VR-based emergency preparedness training from 2010 through 2024. The field's early years saw only scattered publications, for example, some subsequent years (2011, 2013, 2016–2017) had no indexed publications in our dataset. A modest uptick began around the mid-2010s (e.g., 3 studies in 2014 and 4 in 2015), followed by a more pronounced rise after 2018. From 2018 onward, research activity became much more consistent, with multiple publications each year and a clear overall growth trajectory. The output peaked in 2022 with 12 publications, reflecting the highest annual productivity to date. While the counts for 2023 (9 publications) and especially 2024 (4 publications) appear lower, this is likely due to indexing delays for the most recent year rather than an actual decline in scholarly interest. In sum, the trend progresses from a trickle of early studies to sustained engagement by the early 2020s, paralleling the wider adoption of VR technology and an increased recognition of its potential for emergency training.

Top Publishing Venues

Figure 3 presents the leading publication venues for research on VR-based disaster preparedness training for the general public. The distribution of publications is spread across both journals and conference proceedings, with a few venues standing out by hosting multiple studies. Notably, the International Journal of Disaster Risk Reduction, Advanced Engineering Informatics, and Springer's Lecture Notes in Computer Science (LNCS) series each account for three publications in our dataset. Several other outlets have published two papers each, including the ACM International Conference Proceedings Series, the MDPI journal *Applied Sciences*, *Frontiers in Psychology*, and Elsevier's *Fire Safety Journal*. A long list of additional venues, such as specialized IEEE conferences (e.g., ISMAR-Adjunct, TALE) and workshops (e.g., CEUR Workshop Proceedings) appear with a single publication for each.

This venue profile reflects the interdisciplinary nature of the field. Engineering and technology-focused outlets (for instance, the various IEEE-sponsored conferences and transactions, which collectively contribute eight of the 66 studies) feature prominently, underscoring the strong technical and simulation-driven core of this research area. At the same time, the presence of safety science journals (e.g., disaster risk reduction and fire safety venues) and even a psychology journal signals that VR emergency training work spans both technical development and human-centered evaluation. In other words, the field bridges traditional academic research domains and more applied, practice-oriented forums. The balance between diverse journals and conference proceedings suggests that researchers are disseminating findings in both scholarly and practitioner communities, which can help drive technological innovation alongside practical implementation in emergency preparedness training.

Figure 3. Top Publishing Venues (≥ 2 papers) in VR-Based Emergency Preparedness Training (2010–2024).

Keyword Co-occurrence Analysis

To gain insight into prevailing research themes, a keyword co-occurrence analysis was performed on the selected studies (Figure 4). Using VOSviewer ((CRED), 2022), we mapped how often key terms appeared together in titles and author-provided keywords, revealing several core thematic clusters in the field. As expected, foundational terms such as **"virtual reality," "personal training," "e-learning,"** and **"serious games"** form the central hub of the network, these terms are highly prevalent across studies given the focus of our review. This cluster underscores the field's technological and pedagogical foundation, leveraging immersive platforms and game-based learning for disaster preparedness. Another major cluster highlights **emergency scenario and safety drill** themes. Here, keywords such as "evacuation," "fire drill," "hazard," and "preparedness" coalesce, reflecting a strong focus on fire safety and evacuation training. The prominence of *"fire drill"* within this cluster aligns with the many studies dedicated to VR fire evacuation exercises for the general public. A third cluster is oriented around **human–computer interaction (HCI) and user experience**, evidenced by terms such as *"immersion," "usability," "user experience," "realism,"* and (in a few cases) *"augmented reality."* This indicates a thematic focus on the quality of the virtual environment and interface, for example, ensuring realistic immersion and comfort as an important strand of research. Together, these clusters delineate the field's primary concentrations: the development of VR/serious games as training **tools**, the design of realistic **training scenarios** (especially fire emergencies and evacuations), and the enhancement of **user experience** in those virtual environments.

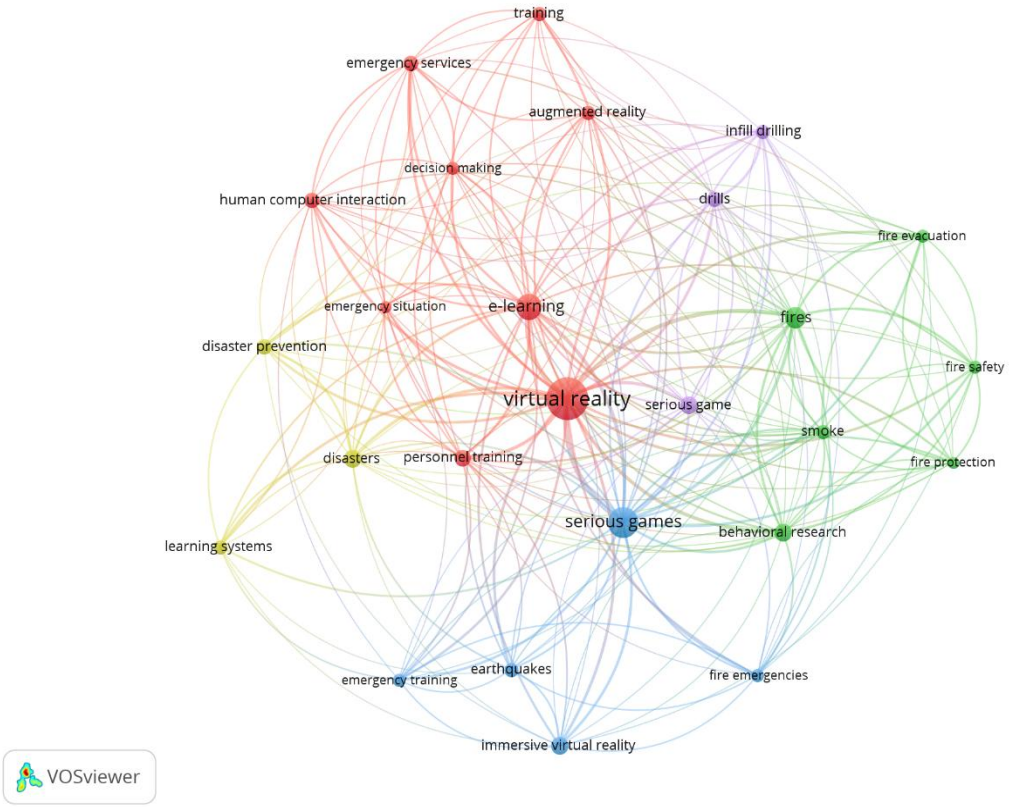

Figure 4. Co-Occurrence Network of Author Keywords in VR-Based Emergency Preparedness Training Literature (2010–2024), Generated Using VOSviewer.

Types of Disasters Addressed in VR Emergency Training

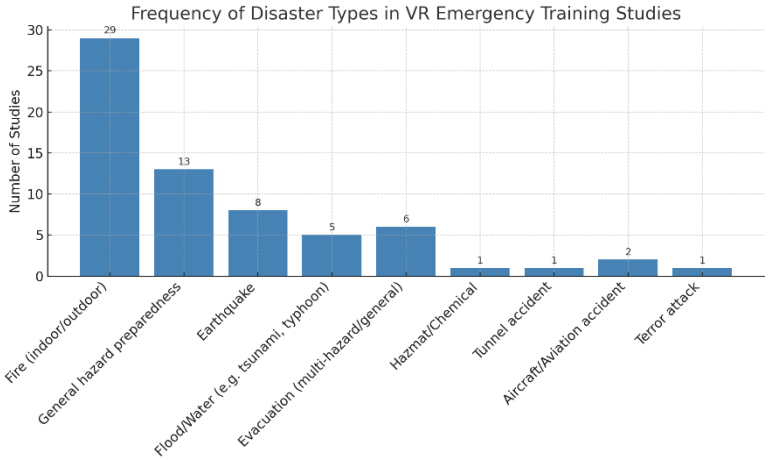

Figure 5. Distribution of Disaster Types Represented in VR-Based Emergency Training Studies (2010–2024).

Analysis of the 66 included studies revealed a strong concentration on certain hazard types in VR-based emergency training research. As shown in *Figure 5*, fire-related emergencies dominate the field (n=29), accounting for the largest proportion of studies. These often focus on *indoor fire* scenarios in built environments such as schools, offices, or residential buildings, with the primary objective of improving evacuation procedures and hazard recognition, though a few studies investigate the wildfires area as well. Earthquakes (n=8) and flood or water-related disasters such as tsunamis and typhoons (n=5) appear less frequently, despite their global relevance. Even less frequently represented are landslides/mudslides, extreme weather events, and specialized hazards such as industrial explosions. Rare or specialized hazards, including chemical spills, tunnel accidents,

aircraft crashes, and terror attacks, each account for only one or two studies, underscoring how unevenly research attention is distributed across different disaster contexts. A smaller but notable number of studies address general hazard preparedness (n=13) and multi-hazard evacuation training (n=6), showing interest in broader, cross-cutting emergency skills.

Evaluation Methods

To assess VR-based emergency training, each study was coded according to the evaluation approaches reported. An inductive coding process was used to group the evaluation methods into the following categories:

- **Self-report (Usability/UX & Presence)** – e.g. questionnaires, surveys, Likert-scale ratings, or standardized user experience instruments (such as System Usability Scale (SUS), User Experience Questionnaire (UEQ), Simulator Sickness Questionnaire (SSQ)).
- **Behavioral Performance** – direct observations or in-application log metrics capturing user actions, decision-making processes, or task completion performance during the simulation.
- **Controlled Comparison** – comparative study designs that evaluate VR training against another condition (for example, comparing outcomes for a VR-trained group vs. a group with traditional training or no training).
- **Knowledge Tests** – pre- or post-training quizzes and knowledge assessments (covering factual or procedural knowledge, often to measure learning gains or retention).
- **Qualitative (Interviews/Focus Groups)** – collection of open-ended feedback from participants (through interviews, focus group discussions, or open-ended survey questions) and subsequent thematic analysis of their responses.
- **Physiological** – biometric measurements during or post-training, such as heart rate, electroencephalography (EEG), or skin conductance, to gauge stress, engagement, or other responses.
- **Long-term Retention/Follow-up** – delayed post-tests or follow-up evaluations days or weeks after the initial training session, aimed at assessing knowledge/skill retention over time.

The analysis revealed a reliance on a few dominant methods, while more robust or long-term approaches remain rare.

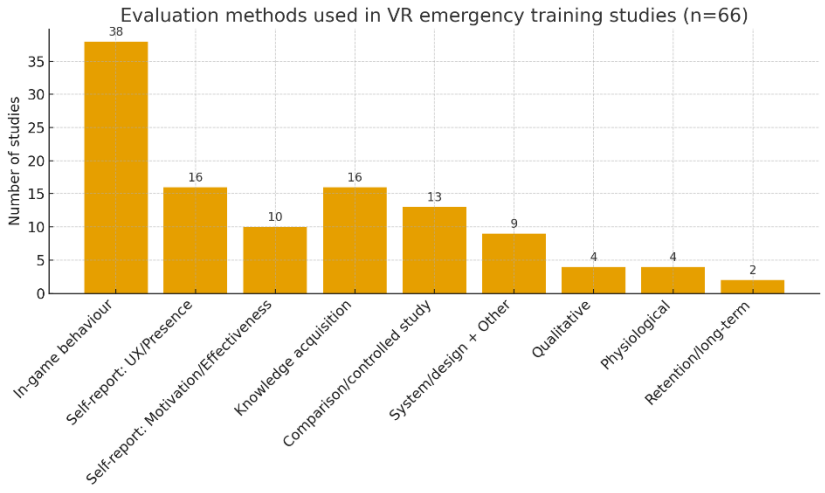

Figure 6. Evaluation Methods Used in VR-Based Emergency Preparedness Training Studies (n = 66).

Figure 6 shows the range of evaluation methods used across the 66 studies. The most common approach was in-game behaviour tracking (n=38), often measuring evacuation time, navigation choices, or other performance indicators directly from the VR environment. A large share of studies also relied on self-report measures, though these appeared in two distinct forms: UX/Presence questionnaires (n=16), focused on immersion, usability, and realism, and perception/motivation questionnaires (n=10), where participants reflected on effectiveness, learning, or motivation. In parallel, knowledge acquisition tests (n=16) and controlled comparisons (n=13) were frequently used to assess learning gains or to contrast VR with traditional methods.

Some papers applied qualitative interviews (n=4), offering open-ended insights into user experience and learning. Only two studies attempted any form of long-term retention testing, underscoring how little is known about whether VR training effects persist over time. A small number of contributions (n=9) focused primarily on

scenario or system design, aiming to model realistic hazards or evaluate the fidelity of VR training environments without involving human performance testing.

Overall, while behavioral tracking is strongly represented, most studies combined multiple approaches, nearly all of which used questionnaires in some form. This reliance on short-term, subjective measures leaves important questions about the durability and transfer of training unanswered. The next section addresses how these methods were combined, showing the mixed-method patterns that dominate the field.

Mixed Methods in Practice

The combined multiple evaluations (37%) are shown below on the bubble matrix. The most frequent pairings were:

- Self-report + Behavioral Performance
- Self-report + Knowledge Tests
- Behavioral Performance + Knowledge Tests

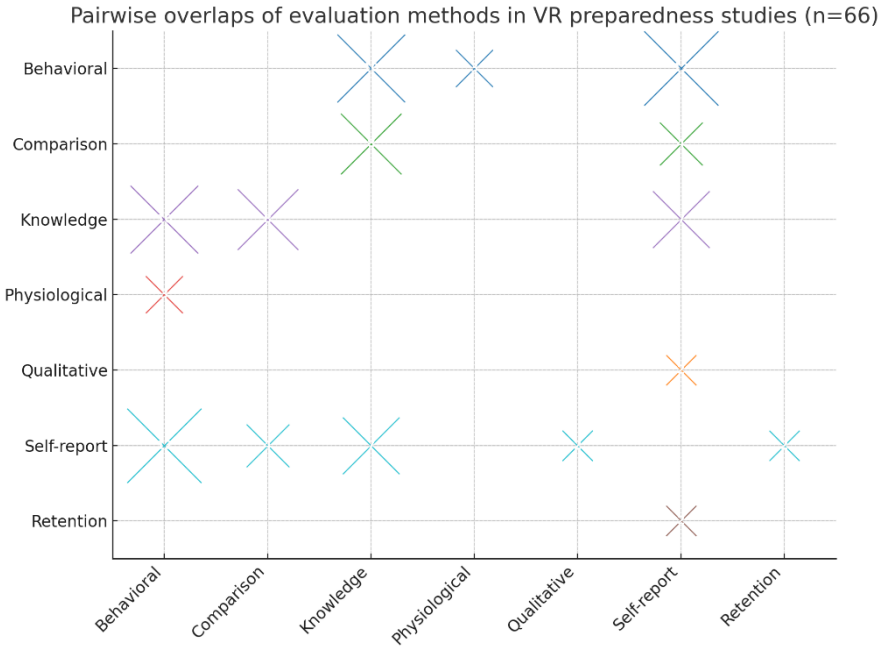

Figure 7: Bubble Matrix Showing the Use of Evaluation Methods in VR-Based Emergency Preparedness Training Studies. Cross-Size Labels Indicate How Many Studies Used Each Pair of Methods in the Same Paper. Rows/Columns List the Seven Categories.

Figure 7 presents a bubble matrix illustrating the relationships between evaluation methods. The matrix shows that mixed evaluations most commonly pair in-game behavioural tracking with self-report questionnaires (n=12) and behavioural measures with knowledge acquisition tests (n=10). These reflect the common practice of pairing observable performance with subjective feedback or pre- and post-learning tests. Smaller but still notable overlaps include behavioral data with physiological measures such as heart rate or eye-tracking (n=3) and controlled-comparison studies with knowledge acquisition (n=8), which represent the most rigorous attempts to benchmark VR training against traditional approaches. Less frequent pairings involve self-reports with comparison studies (n=4), self-reports with qualitative interviews (n=2), and rare long-term retention follow-ups (n=2).

Overall, the figure highlights that while most studies combine methods, these combinations are clustered around a few pairings. Only two studies include any form of follow-up post-test, highlighting how little is known about whether learning from VR persists beyond the immediate session. The limited use of long-term or cross-scenario assessments reflects practical constraints, real disasters cannot be recreated, but it also leaves uncertainty about retention and transfer. In practice, “mixed methods” in this field typically means combining subjective feedback with one short objective measure. This explains why many studies report strong engagement and immediate task performance, yet provide limited evidence of durable learning or superiority over non-VR training.

Evaluating the long-term effectiveness of disaster-preparedness training through VR remains inherently difficult, as real emergencies occur rarely and under unpredictable conditions. Consequently, manual knowledge assessments and delayed follow-up tests remain the most practical ways to estimate learning retention and compare VR with traditional training methods such as web-based or video materials.

Key Findings, Implications, and Future Directions

The bibliometric analysis reveals several structural characteristics of the current research landscape on VR-based emergency preparedness training for civilians. Across the 66 studies included in this review, research activity is strongly concentrated on short-term, scenario-based training applications, most commonly focusing on evacuation and hazard recognition in fire, earthquake, and flood contexts. Fire-related scenarios dominate the dataset (n=29), while other hazards such as floods (n=5), earthquakes (n=8), and other disaster contexts appear far less frequently. This distribution suggests that research attention has concentrated on hazards that are frequently addressed in safety training contexts, particularly building fires and evacuation scenarios, while other hazards, such as floods or tsunamis appear less frequently in the literature.

A second notable pattern concerns how VR training interventions are evaluated. Most studies rely on short-term measurements immediately after the training session. Behavioural performance tracking within the VR environment is the most frequently used approach (n=38), often measuring navigation choices, evacuation time, or task completion. Self-report questionnaires related to usability, presence, or perceived learning are also widely used (UX/presence measures n=16; perception or motivation questionnaires n=10). Knowledge-based assessments are present in 16 studies, typically using pre- and post-training quizzes to estimate learning gains.

However, more rigorous evaluation approaches remain relatively rare. Only two of the 66 studies included delayed follow-up assessments designed to examine long-term retention of knowledge or skills. Similarly, physiological measures such as heart rate or EEG were used in only a small subset of studies (n=3), indicating that objective measurement of cognitive or emotional responses during VR training remains uncommon in this research domain. These patterns suggest that while immersive VR environments enable detailed behavioural tracking, most studies still rely primarily on short-term, often subjective indicators of training outcomes.

Another characteristic emerging from the dataset is the predominance of single-session training experiences. Most interventions are implemented as isolated simulation sessions rather than repeated training programs or longitudinal preparedness initiatives. As a result, the literature provides limited evidence regarding whether skills practiced in VR are retained over time or transferred to real-world preparedness behaviour.

Beyond methodological aspects, the bibliometric results also highlight thematic concentration within the field. Keyword co-occurrence analysis shows that the research landscape clusters around three main themes: immersive training technologies (virtual reality and serious games), emergency scenario simulation (particularly evacuation and fire drills), and user experience or human–computer interaction aspects such as immersion, realism, and usability. These clusters indicate that current research simultaneously addresses technological development, scenario design, and user experience considerations.

Taken together, these patterns suggest that the field of VR-based emergency preparedness for civilians is still in a formative stage. Research output has grown rapidly since 2018, but methodological diversity and hazard coverage remain uneven. Future work may benefit from expanding the range of disaster scenarios investigated, incorporating longitudinal or repeated-training designs, and exploring evaluation approaches that connect virtual performance with broader preparedness outcomes.

DISCUSSIONS AND CONCLUSION

This bibliometric review mapped the development of research on VR-based disaster preparedness training for civilians between 2010 and 2024. The analysis shows a clear growth of publications after 2018, reflecting the broader accessibility of consumer VR technologies and increasing interest in immersive training tools. However, the distribution of topics and evaluation approaches indicates that the field remains uneven in both thematic coverage and methodological depth.

Several recent review studies have examined the use of immersive VR technologies for safety and emergency preparedness training (Alshowair et al., 2024; Stefan, Mortimer, & Horan, 2023; Strojny & Dużmańska-Misiarczyk, 2023). These studies review VR-based training applications and discuss evaluation approaches for assessing training outcomes. Although their scope and inclusion criteria differ from those of the present study, the overall trends reported in those reviews are broadly consistent with the patterns observed in our dataset. In

particular, both the prior literature and the present bibliometric mapping highlight growing research interest in immersive preparedness training and different evaluation approaches.

One notable pattern is the concentration of research around a limited set of disaster scenarios, particularly fire evacuation training. Expanding the diversity of simulated hazards may therefore represent an important direction for future research.

Another important observation concerns how VR-based preparedness interventions are assessed. A fundamental challenge in this domain lies in linking virtual reality training to real-world disaster behaviour. Actual disaster events are unpredictable, rare, and cannot be staged under controlled conditions with genuine risk. As a result, evidence that VR-trained participants perform better under real emergency pressure remains limited (Stefan, Mortimer, & Horan, 2023). Longitudinal or naturalistic follow-up studies that track how individuals behave months or years after VR training are virtually absent. Such studies would require both institutional cooperation and large-scale adoption, which the field has not yet achieved. Consequently, while VR offers an immersive approximation of hazardous events, it cannot fully replicate the cognitive, emotional, and social complexity of real-world disasters.

In accordance with earlier studies, this study acknowledges that current VR captures public attention more effectively than static text or video-based materials and can act as a powerful entry point into preparedness education. However, current simulations are typically pre-scripted and hazard-specific, meaning they cannot represent the full variability of real disasters. Floods, fires, and earthquakes differ not only in their physical dynamics but also in their decision-making demands, stress levels, and environmental cues (Lovreglio, 2020). VR, therefore, tends to train responders firstly within bounded scenarios and only after supporting broader adaptive preparedness.

This imbalance highlights opportunities for future research to diversify scenario design, integrate underrepresented hazard types, and explore whether training in one hazard domain transfers to others. Scenario design could also benefit from community-level data and historical disaster records to simulate more complex and realistic environments (Keya et al., 2024). Emerging VR platforms that integrate real-time environmental data may help bridge the gap between simulated and real-world preparedness contexts, although such validation remains challenging. Moreover, collaborative and multiplayer VR scenarios could simulate collective evacuation, communication, and coordination under stress, skills that individual simulations cannot address.

Future research should also address accessibility, digital access, and usability for diverse and vulnerable populations, ensuring that innovations in VR do not widen the digital divide. Opportunities also exist to explore collaborative and social learning approaches that leverage information about community dynamics to strengthen preparedness.

This review has several limitations. The dataset was limited to English-language publications indexed in major academic databases, and some local training initiatives or grey literature may therefore be underrepresented. In addition, evaluation categories were coded primarily from titles and abstracts, and publication counts for 2024 may be affected by indexing lag. These limitations do not alter the general patterns observed but highlight the need for continued monitoring of developments in this rapidly evolving field.

In summary, VR currently functions most effectively as an awareness and engagement tool that can complement existing preparedness education approaches. For the technology to mature as a training medium, scenario design must increasingly incorporate important realistic environmental data and community contexts, while future studies should include: defining what realistic data means in a context, and how to follow-up evaluations, and perform repeated-training assessments for knowledge transfer. Through further methodological advancement, VR may evolve from an experimental awareness tool into a more validated component of public safety education, supporting both engagement toward retained knowledge, skill transfer, and real-world performance.

ACKNOWLEDGEMENT

We would like to thank Gøril Tvedten Jorem for her guidance regarding the literature search process and the project B-prepared for funding. B-prepared is co-funded by the European Commission under the Horizon Europe Programme (Grant ID: 101121134).

References

(CRED), C. f. R. o. t. E. o. D. (2022). *Annual Disaster Statistical Review 2022: The numbers and trends*. https://www.cred.be/sites/default/files/2022_EMDAT_report.pdf

Abbas, J. R., Chu, M. M. H., Jeyarajah, C., Isba, R., Payton, A., McGrath, B., Tolley, N., & Bruce, I. (2023). Virtual reality in simulation-based emergency skills training: A systematic review with a narrative synthesis. *Resuscitation Plus*, *16*, 100484. https://doi.org/10.1016/j.resplu.2023.100484

Alshowair, A., Bail, J., AlSuwailem, F., Mostafa, A., & Abdel-Azeem, A. (2024). Use of virtual reality exercises in disaster preparedness training: A scoping review. *SAGE Open Medicine*, *12*, 20503121241241936. https://doi.org/10.1177/20503121241241936

Buxton, B., & Fitzmaurice, G. W. (1998). HMDs, caves & chameleon: a human-centric analysis of interaction in virtual space. *SIGGRAPH Comput. Graph.*, *32*(4), 69–74. https://doi.org/10.1145/307710.307732

Cha, M., Han, S., Lee, J., & Choi, B. (2012). A virtual reality based fire training simulator integrated with fire dynamics data. *Fire safety journal*, *50*, 12-24.

Chittaro, L., & Buttussi, F. (2015). Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety. *IEEE transactions on visualization and computer graphics*, *21*(4), 529-538.

del Carmen Cardós-Alonso, M., Otero-Varela, L., Redondo, M., Uzuriaga, M., González, M., Vazquez, T., Blanco, A., Espinosa, S., & Cintora-Sanz, A. M. (2024). Extended reality training for mass casualty incidents: a systematic review on effectiveness and experience of medical first responders. *International Journal of Emergency Medicine*, *17*(1), 99. https://doi.org/10.1186/s12245-024-00685-3

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of business research*, *133*, 285-296.

Feng, Z., González, V. A., Amor, R., Lovreglio, R., & Cabrera-Guerrero, G. (2018). Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Computers & Education*, *127*, 252-266. https://doi.org/10.1016/j.compedu.2018.09.002

Feng, Z., González, V. A., Mutch, C., Amor, R., & Cabrera-Guerrero, G. (2021). Instructional mechanisms in immersive virtual reality serious games: Earthquake emergency training for children. *Journal of Computer Assisted Learning*, *37*(2), 542-556.

Heldal, I., Wijkmark, C. H., & Pareto, L. (2016). Simulation and serious games for firefighter training: Challenges for effective use. *Nokobit*, *24*(1), 1-12.

Kanade, S. G., & Duffy, V. G. (2022). Use of virtual reality for safety training: a systematic review. In *International Conference on Human-Computer Interaction* (pp. 364-375). Springer. https://doi.org/10.1007/978-3-031-06018-2_25

Kaplan, A. D., Cruit, J., Endsley, M., Beers, S. M., Sawyer, B. D., & Hancock, P. (2020). The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: a meta-analysis. *Human factors*, 0018720820904229.

Keya, R. T., Heldal, I., Frgland, T. H., Székely-Keresztesi, O., Székely, Z., & Kon, E. (2024). Emergency Preparedness Training Using Virtual Reality Games: Allowing Knowledge Transfer in the Digital Age. In *European Conference on Knowledge Management* (pp. 369-378). Academic Conferences International Limited. https://doi.org/10.34190/eckm.25.1.2890

Keya, R. T., Heldal, I., Patel, D., Murano, P., & Wijkmark, C. H. (2025). Implementing Virtual Reality for Fire Evacuation Preparedness at Schools. *Computers*, *14*(7), 286.

LaValle, S. M., Yershova, A., Katsev, M., & Antonov, M. (2014). Head tracking for the Oculus Rift. 2014 IEEE international conference on robotics and automation (ICRA),

Lochmannová, A., Šimon, M., Hořejší, P., Bárdy, M., Reichertová, S., & Gillernová, K. (2022). The use of virtual reality in training paramedics for a mass casualty incident. *Applied Sciences*, *12*(22), 11740.

Lovreglio, R. (2020). Virtual and Augmented reality for human behaviour in disasters: a review. Fire and Evacuation Modeling Technical Conference (FEMTC) 2020 Online Conference,

Paton, D. (2013). Disaster resilient communities: developing and testing an all-hazards theory. *IDRiM Journal*, *3*(1), 1-17.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778.

Scorgie, D., Feng, Z., Paes, D., Parisi, F., Yiu, T. W., & Lovreglio, R. (2024). Virtual reality for safety training: A systematic literature review and meta-analysis. *Safety Science*, *171*, 106372. https://doi.org/10.1016/j.ssci.2023.106372

Shaw, E., Roper, T., Nilsson, T., Lawson, G., Cobb, S. V., & Miller, D. (2019). The heat is on: Exploring user behaviour in a multisensory virtual environment for fire evacuation. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*.

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing Our Lives with Immersive Virtual Reality [Opinion]. *Frontiers in Robotics and AI*, *Volume 3 - 2016*. https://doi.org/10.3389/frobt.2016.00074

Stefan, H., Mortimer, M., & Horan, B. (2023). Evaluating the effectiveness of virtual reality for safety-relevant training: a systematic review. *Virtual Reality*, *27*(4), 2839-2869.

Strojny, P., & Dużmańska-Misiarczyk, N. (2023). Measuring the effectiveness of virtual training: A systematic review. *Computers & Education: X Reality*, *2*, 100006.

Susi, T., Johannesson, M., & Backlund, P. (2007). *Serious games: An overview*. https://api.semanticscholar.org/CorpusID:13561855

Székely, Z., Székely-Keresztesi, O., Márkus, Z. L., Majdik, A., & Szkaliczki, T. (2025). The Role of Science, Technology, Innovation, Education and Gaming in Building a Culture of Disaster Preparedness with the B-prepared Project. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, *15*, 49-54.

Tang, Y.-M., Au, K. M., Lau, H. C., Ho, G. T., & Wu, C.-H. (2020). Evaluating the effectiveness of learning design with mixed reality (MR) in higher education. *Virtual Reality*, *24*(4), 797-807.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., & Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, *169*(7), 467-473. https://doi.org/10.7326/M18-0850

van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523-538. https://doi.org/10.1007/s11192-009-0146-3

Wijkmark, C. H. (2024). *Virtual reality for fire and rescue service professionals' training* [Doctoral dissertation, Western Norway University of Applied Sciences].

Williams-Bell, F. M., Murphy, B. M., Kapralos, B., Hogue, A., & Weckman, E. J. (2015). Using Serious Games and Virtual Simulation for Training in the Fire Service: A Review. *Fire Technology*, *51*(3), 553-584. https://doi.org/10.1007/s10694-014-0398-1

Zhu, Y., & Li, N. (2021). Virtual and augmented reality technologies for emergency management in the built environments: A state-of-the-art review. *Journal of Safety Science and Resilience*, *2*(1), 1-10.