

# Interactive 3D Digital Environments as Ergonomic Data Collection Platforms: A Cross-Immersion Analysis

Huishu Deng, Li Zhang, and Xinyue Wang

School of Architecture, Tsinghua University, Beijing, China

## ABSTRACT

Interactive 3D digital environments have evolved from mere representation media into multifunctional platforms for collecting granular ergonomic and behavioral data. While previous reviews established a foundation for behavior mapping in immersive virtual environments, the rapid emergence of diverse interaction modalities and the integration of multimodal sensing in recent years necessitate a systematic update. This study reviews 208 journal papers from the Web of Science Core Collection (2023–2026) and employs a large language model (DeepSeek V4-Pro) for structured metadata extraction. The analysis compares five data collection categories (physiological, spatiotemporal, foreground interaction, subjective feedback, and contextual data) and eight ergonomic issue classes across four immersion modalities: mobile apps, screen-based open worlds, augmented reality (AR), and immersive virtual reality (VR). The results indicate that immersive VR remains the dominant modality (75.1%) for high-fidelity physiological measurement and cognitive workload analysis, particularly through integrated eye tracking. Screen-based open worlds demonstrate relative efficiency in spatial behavior tracking, while AR presents a balanced profile between interaction usability and physical risk assessment. Mobile apps, by contrast, remain underexplored with only sporadic publications. The cross-modality comparison of ergonomic issues further reveals a clear alignment between immersion level and research focus: lower-immersion modalities concentrate on spatial behavior and HCI, whereas higher-immersion platforms enable comprehensive physiological and cognitive monitoring. Meanwhile, this study demonstrates the feasibility of LLM-assisted information extraction for longitudinal ergonomic research. The findings offer a decision-making guide for selecting appropriate interactive 3D platforms based on specific ergonomic research objectives.

**Keywords** 3D digital environments, Virtual reality, Ergonomic data, Human-space interaction, Large language model

## INTRODUCTION

With the rapid development of digital media, game engines, wearable sensors, and immersive visualization technologies, interactive 3D digital environments have gradually evolved from visualization tools into experimental platforms for behavioral and ergonomic research. Compared with traditional questionnaires or direct observation, interactive 3D environments provide

controllable, repeatable, and data-rich experimental conditions, enabling researchers to capture spatial behavior, interaction trajectories, cognitive responses, and physiological signals simultaneously. Such platforms are increasingly applied in architecture, urban studies, and human-computer interaction (HCI) to evaluate human-space synergy.

The utility of immersive technologies has been rigorously scrutinized. Lanier et al. (2019) highlighted the need for experimental standards by questioning Virtual Reality (VR) research validity, while Kalantari et al. (2024) validated VR as a reliable proxy by demonstrating navigation consistency between virtual and real environments. In safety contexts, Kim et al. (2019) and Stefan et al. (2023) proved the effectiveness of Augmented Reality (AR) and VR in enhancing training outcomes compared to traditional methods.

Integration of physiological metrics has further deepened user analysis. Kim and Lee (2021) and Zheleva et al. (2023) utilized eye tracking and psycho-physiological measures to calibrate and optimize virtual experiences. Crucially, Kumar and Dhar (2023) provided a comprehensive mapping review up to early 2023, establishing a baseline for human-building interaction while noting fragmentation across modalities. Looking ahead, Avilés-Castillo et al. (2025) and Krupas et al. (2024) have explored evolving user experiences and human-centric digital twins for collaboration.

Expanding on the foundation of Kumar and Dhar (2023), this study conducts a systematic review of interactive 3D digital environments for ergonomic data collection from 2023 to 2026. This research evaluates the correspondence between diverse immersion levels and their ergonomic data acquisition capacities. By systematically analyzing technological characteristics across multiple modalities, we identify prevailing trends and critical research gaps in the integration of 3D digital platforms. In addition, this study explores the feasibility of using large language models to support structured literature extraction and comparative analysis.

## RESEARCH METHODOLOGY

This study employed a combined approach of systematic search, large language model assisted information extraction, and normalized multi-dimensional comparison.

### Literature Search Strategy

This study used the Web of Science Core Collection as the literature source, with the search period limited to 2023–2026, document types restricted to Article and Proceedings Paper, and the language restricted to English. A search query was constructed around three conceptual dimensions: environments, ergonomic data, and objectives. Within each dimension, keywords were connected with the OR operator, and the three dimensions were connected with the AND operator. The specific keywords are listed in Table 1.

**Table 1:** Keyword categories for literature search.

Dimension	Environments	Ergonomic Data	Objectives
Search Terms	3D digital environment	behavioral data	data collection
	virtual environment	cognitive data	data acquisition
	virtual reality	wayfinding	measurement
	augmented reality	eye tracking	assessment
	screen-based	spatial behavior	data logging
	desktop	movement trajectory	behavioral logging
	non-immersive	gaze	interaction log
	mobile game	dwelling time	in-game questionnaire
	mobile app	clickstream	embedded survey
	touch screen	click heatmap	usability evaluation
	touchscreen	touch interaction	ergonomic
	tablet-based	trajectory data	
	smartphone-based	movement pattern	
	open world	spatial cognition	
	sandbox game	cognitive map	
	serious game	route choice	
	video game	preference data	
	interactive 3D	task completion time	
	web-based 3D		
	panoramic tour		

### Information Extraction Using a Large Language Model

For the 332 retained papers, the application programming interface of the large language model DeepSeek V4-Pro was employed for structured information extraction. The first 15 pages of each PDF were converted to plain text and submitted to a predefined system prompt. The prompt required the model to output a strict JSON object with fields covering the following dimensions: bibliographic metadata (title, year, authors, journal, research field), environmental attributes (environment type, immersion level, sample size, participant type), data collection characteristics (data categories and specific metrics, e.g., spatiotemporal trajectories, foreground interactions, subjective feedback, physiological signals, contextual data), ergonomic concerns (research questions, human–computer interaction mode, human-space interaction cognitive aspects), spatial attributes (spatial scale, specific spatial scene), and statements related to ecological validity. In addition, the model was required to explicitly determine whether each paper described an interactive 3D digital environment and to provide a rationale.

During extraction, JSON validation and exception handling were applied to automatically skip illformatted entries, and a checkpointbased resumption strategy was adopted to guarantee data integrity. To verify the reliability of the large language model extraction results, 20 papers were randomly selected from the included literature for manual verification. The reviewer crosschecked the full text against the model outputs, examining key fields such as the title, year, environment type, and data collection categories item by item, and confirmed that the extracted information was consistent with the original text, with no notable errors identified.

## Data Screening and Multidimensional Comparative Analysis

After information extraction, 208 papers classified as interactive 3D digital environments were retained and grouped into four modality categories: mobile apps, screen-based open worlds, augmented reality (AR), and immersive virtual reality (VR). The analysis comprised two parts. The first part was statistical description: the frequency and proportion of environment types, data collection categories, ergonomic issue types, and spatial scales were calculated by publication year. The second part was cross-classification comparison: using modality as the grouping variable, the distribution of data collection types was compared through normalized proportions; for ergonomic issue categories, weighted normalization was applied to account for sample size disparities across modalities. The proportion of papers discussing ecological validity was also examined for each modality. These normalized comparisons aimed to reveal the correspondence between immersion level and the data collection capacity of 3D digital environments.

## RESULTS AND DISCUSSION

### Descriptive Statistics

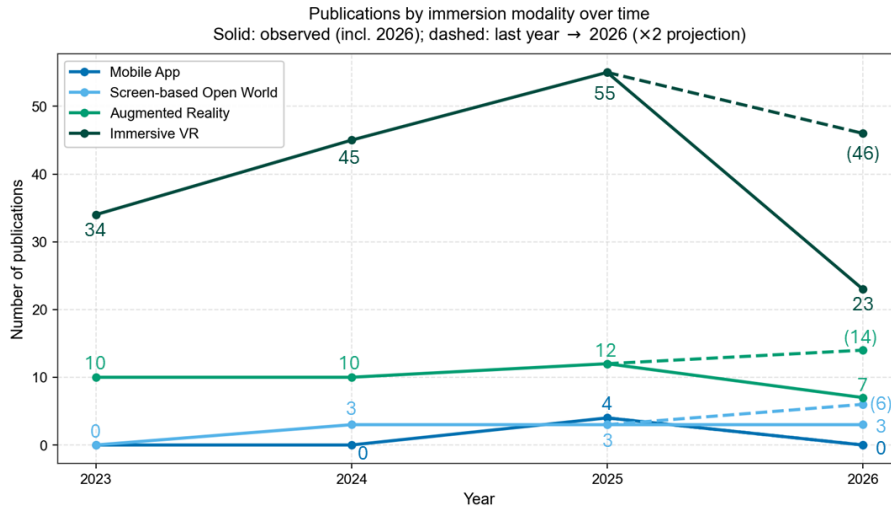
#### Annual Publication Trends Across Immersion Modalities

Among the 208 papers categorized into the four immersion modality groups, the annual distribution of publications for mobile apps, screen-based open worlds, augmented reality, and immersive virtual reality is shown in Figure 1. It should be noted that since the literature search was conducted up to mid-2026, the statistical values for this year were estimated by doubling the actual observed counts from the first half of the year, in order to maintain continuity of the trend lines in the figure. Accordingly, the projected full-year values for 2026 are indicated with dashed lines, while the actual observed values for the first half of the year are plotted with solid lines.

Looking at the four-year totals, immersive VR led with 157 papers, accounting for 75.1% of all modality-classified papers. Augmented reality ranked second with 39 papers (18.7%), while screen-based open worlds and mobile apps accounted for only 9 (4.3%) and 4 (1.9%) papers, respectively. Together, these two low-immersion modalities constituted less than 7% of the total, indicating that current research in this area remains heavily reliant on head-mounted immersive display devices, while screen-based and mobile platforms with lower immersion levels have yet to be fully explored.

In terms of annual trends, immersive VR publications rose from 34 in 2023 to 45 in 2024 and further to 55 in 2025, before the estimated value declined to 46 in 2026. This rise-and-fall pattern may be influenced by multiple factors: the early growth likely reflects the progressive maturation of virtual reality experimental paradigms, while the recent dip may be associated with a partial shift of research interest toward augmented reality, and is also subject to uncertainty given that the 2026 figure is an estimate. Augmented reality showed a steady, modest increase, from 10 papers in 2023 to 12 in 2025, with an estimated 14 in 2026, suggesting a slowly growing appeal of augmented reality for ergonomic data collection. Screen-based open worlds

remained stable at 3 papers in both 2024 and 2025, with an estimated 6 in 2026. Although the absolute number is small, the data suggest an emerging upward trend. Mobile apps had no related publications in 2023 or 2024, with 4 papers appearing in 2025, indicating that research in this modality is still at a sporadic, exploratory stage.



**Figure 1:** Publications by immersion modality over time.

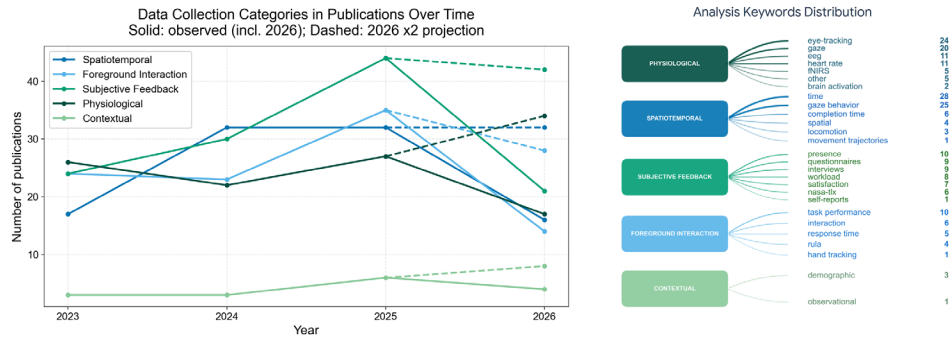
In summary, research activity varies considerably across immersion modalities. Immersive VR continues to dominate, but the growth signals observed in augmented reality and screen-based open worlds may suggest that behavioral data collection in lower-immersion environments is beginning to form a new research branch.

### Frequency and Annual Trends of the Five Data Collection Categories

Among the 208 papers, the occurrence frequencies of the five data collection categories were: physiological 172 (82.3%), spatiotemporal 149 (71.3%), foreground interaction 127 (60.8%), subjective feedback 112 (53.6%), and contextual 35 (16.7%). Most papers collected two or more categories, reflecting a multimodal fusion trend. High-frequency keywords within each category showed that physiological data were dominated by eye tracking and gaze, spatiotemporal data by time metrics, gaze behavior, and completion time, subjective feedback by presence, questionnaires, and workload, foreground interaction by task performance, response time, and RULA, and contextual data by demographics and observational records (Figure 2, right).

The annual trends of the five data categories showed clear divergence (Figure 2, left). Physiological data maintained a continuous rise and is expected to keep growing; foreground interaction and subjective feedback data are projected to decline after peaking in 2025; spatiotemporal data rose then fell; and contextual data remained low and stable. These findings indicate that physiological measurement, especially eye tracking as a cross-modality tool,

has become the core collection approach, while the earlier steady growth of subjective feedback data reflects increasing demand for quantifying user experience, and multimodal fusion has become the prevailing paradigm.



**Figure 2:** Annual distribution of publications across the five data collection categories.

### Ergonomic Issues Classification

Keyword-type ergonomic issue items were extracted from each paper. For example, Zheleva et al. (2023) yielded items such as “VR psychophysiological calibration,” “inter-individual biometric differences,” “adaptive training,” and “cognitive load measurement”; Kalantari et al. (2024) yielded items such as “ecological validity of VR,” “virtual vs real navigation performance,” “age effects on spatial navigation,” and “wayfinding uncertainty in buildings.” After deduplication and cleaning, a total of 673 valid ergonomic issues were obtained and classified into eight predefined categories. The categories and their representative issues are as follows:

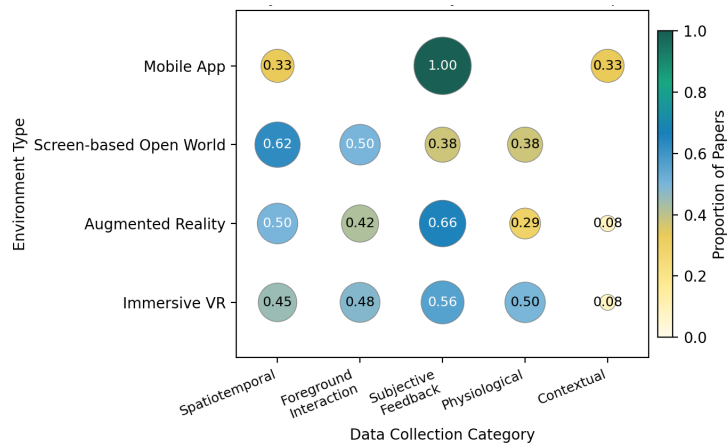
- **Physiological Measurement and Biosignals** was the most prominent category, with typical issues addressing eye tracking, EEG and heart rate measurement, gaze behavior analysis, pupillometry, and fNIRS brain imaging.
- **Human–Computer Interaction and Usability** covered user experience evaluation in virtual environments, interaction mode comparison, presence measurement, interface design, and usability testing, such as VR input modality comparison and AR-HUD usability assessment.
- **Cognitive Workload and Stress** commonly appeared in studies on mental workload measurement in VR cognitive tasks, vigilance monitoring in driving contexts, music performance anxiety assessment, and arachnophobia severity classification.
- **Spatial Behavior and Navigation** focused on wayfinding performance evaluation, route choice modeling, spatial memory testing, building evacuation analysis, and visual attention to signage.
- **Physical Risk and Biomechanical Load** included RULA/REBA posture scoring, manufacturing work posture analysis, exoskeleton fit optimization, spinal force prediction, and repetitive strain risk assessment.
- **Safety and Hazard Detection** was frequently investigated in VR fire evacuation studies, construction site hazard recognition, pedestrian crossing risk assessment, and struck-by accident prevention.

- **Training and Rehabilitation** encompassed immersive dental posture training, joint attention intervention for children with autism, VR cataract surgery simulation validation, cognitive stimulation therapy comparison, and VR-based cognitive rehabilitation for mild cognitive impairment.
- **Design and Evaluation Methodology** addressed virtual prototype assessment, VR ergonomic evaluation tool development, participatory workplace design, and noninferiority trials of remote AR medical simulation.

## Cross-Classification Comparison

### Data Collection Categories Across Modalities

Figure 3 uses a bubble heatmap to display the normalized frequency distribution of five data collection categories across four environment types, with both bubble area and color intensity representing the normalized proportion. Overall, the four modalities exhibited both common patterns and clear differentiation in data collection categories. Physiological signal measurement maintained a high proportion across most modalities, particularly in immersive VR, further confirming its status as the core approach for behavioral data collection in current 3D environments. Spatiotemporal data were also widely used across modalities, indicating that movement trajectories and location recording constitute a fundamental collection dimension shared across platforms.



**Figure 3:** Normalized frequency of data collection categories by immersion modality.

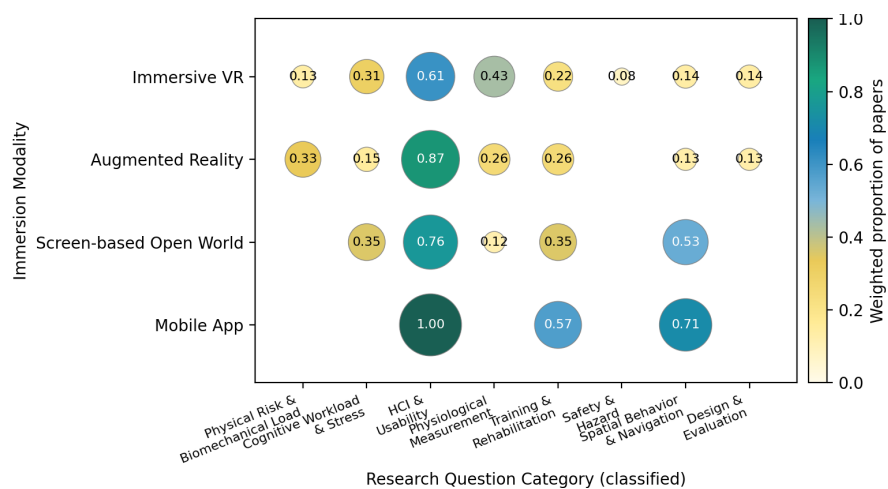
Regarding cross-modality differences, immersive VR showed the most pronounced reliance on physiological signal measurement, with its proportion ranking highest among the four modalities. This is closely related to the hardware advantage of VR headsets, which can seamlessly integrate biosensors such as eye trackers, heart rate monitors, and EEG, making them naturally suited for synchronous multimodal physiological signal acquisition. Screen-based open worlds showed a relatively higher proportion in spatiotemporal data, reflecting the ease of recording continuous

movement trajectories and path choices in that modality. Augmented reality stood out in its proportion of subjective feedback data, possibly due to the frequent need for user evaluation and usability feedback in AR application scenarios. Although the number of mobile app studies was extremely limited, their data collection pattern centered on spatiotemporal data and subjective feedback, consistent with the typical characteristics of lightweight mobile applications.

Foreground interaction data showed a relatively even distribution across the four modalities, suggesting that task performance and interaction behavior measurement serve as cross-modality evaluation tools. Contextual data remained at a low proportion across all modalities, indicating that demographic information and observational records are included only as supplementary background information in a small number of studies and have not yet become a mainstream collection objective.

### Distribution of Ergonomic Issues Across Modalities

Figure 4 presents the weighted proportional distribution of eight ergonomic issue categories across the four modality groups. Overall, human–computer interaction and usability were the only categories that ranked highest or second-highest in all four modalities, with proportions reaching 1.00 for mobile apps, 0.87 for augmented reality, 0.76 for screen-based open worlds, and 0.61 for immersive VR, indicating that regardless of immersion level, interaction design and usability evaluation remain the most prevalent ergonomic concerns in 3D digital environments.



**Figure 4:** Normalized frequency of ergonomic issue categories by immersion modality.

In terms of modality-specific patterns, mobile apps, aside from HCI, showed a proportion of 0.71 in spatial behavior and navigation and 0.57 in training and rehabilitation, while physical risk, cognitive workload, and physiological measurement recorded zero studies, reflecting a current focus entirely on spatial cognition and skill training. Screen-based open worlds recorded 0.53 in spatial behavior and navigation, 0.35 each in cognitive

workload and stress and training and rehabilitation, and only 0.12 in physiological measurement, with physical risk and safety detection at zero, presenting a profile centered on spatial cognition and cognitive assessment with limited attention to physiological and physical concerns.

Augmented reality displayed a relatively balanced spectrum. Beyond HCI, physical risk and biomechanical load accounted for the highest share at 0.33, with physiological measurement and training and rehabilitation each at 0.26, and spatial behavior and design evaluation each at 0.13, while safety detection was absent. This distribution suggests that AR applications, while attending to interaction experience, have directed considerable attention to posture assessment and musculoskeletal risk related to on-site operations, with distinct needs in wearable comfort and field intervention. Immersive VR showed the broadest coverage, with research spanning seven of the eight categories. Physiological measurement and biosignals recorded the highest proportion at 0.43, followed by cognitive workload and stress at 0.31, training and rehabilitation at 0.22, spatial behavior and design evaluation each at 0.14, physical risk at 0.13, and safety and hazard detection at 0.08, making it the only modality covering the safety detection category. This profile confirms the comprehensive capacity of VR in multimodal physiological data acquisition and cognitive state monitoring.

In summary, the ergonomic issue profiles show clear differentiation across modalities: mobile apps focus on spatial behavior and interaction, screen-based open worlds emphasize spatial cognition and cognitive assessment, AR strikes a balance between interaction and physical risk, and VR centers on physiological measurement and cognitive workload while accommodating multidimensional ergonomic evaluation. This distribution pattern also reveals an alignment between immersion level and ergonomic issue type—lower-immersion modalities can already effectively support the collection of spatial behavior data such as wayfinding, whereas higher-immersion modalities offer irreplaceable advantages in physiological signal measurement.

## CONCLUSION

This study systematically reviewed 208 papers (2023–2026) on interactive 3D digital environments as ergonomic data collection platforms, comparing four immersion modalities: mobile apps, screen-based open worlds, augmented reality (AR), and immersive virtual reality (VR). The findings confirm a clear alignment between immersion level and data collection capacity. Immersive VR dominates the current landscape and offers unmatched advantages in high-fidelity physiological measurement and cognitive workload analysis, making it indispensable for research requiring synchronous biosignal acquisition. AR occupies a middle ground, demonstrating a balanced capacity for interaction usability evaluation and physical risk assessment, which aligns with its on-site operational applications. Screen-based open worlds provide efficient and accessible tools for spatial behavior tracking and navigation studies, whereas mobile apps remain underexplored, with only sporadic publications to date, suggesting that their potential for lightweight, large-scale behavioral logging has yet to be realized. The study further indicates that multimodal data fusion has become a major trend, combining

physiological signals, spatiotemporal trajectories, foreground interactions, and subjective feedback to better understand human-space interaction. In addition, the research demonstrates the feasibility of using large language models to support systematic literature analysis and structured information extraction in ergonomic research.

In summary, interactive 3D digital environments are evolving from visualization tools into differentiated ergonomic research infrastructure. Rather than treating immersion as a linear hierarchy, researchers should match platform selection to specific ergonomic objectives. Future research would benefit from standardized cross-modal validation frameworks and expanded exploitation of underrepresented platforms.

## ACKNOWLEDGMENT

This study is funded by the National Natural Science Foundation of China (Grant No. 52278023).

## REFERENCES

- Avilés-Castillo, Fátima. Buele, Jorge. and Palacios-Navarro, Guillermo. (2025). Virtual reality and user experience: Current trends and future challenges, IEEE Access.
- Kalantari, Saleh. Mostafavi, Armin. Xu, Tong Bill. Lee, Anne Seoyoung. and Yang, Qi. (2024). Comparing spatial navigation in a virtual environment vs. an identical real environment across the adult lifespan, *Computers in Human Behavior*, Volume 157.
- Kim, Kyungki. Alshair, Mohammed. Holtkamp, Brian. Yun, Chang. Khalafi, SeyedAmirhesam. Song, Linguang. and Suh, Min Jae. (2019). Using immersive augmented reality to assess the effectiveness of construction safety training, *Journal of Construction Engineering and Project Management*, Volume 9, No. 4. pp. 16–33.
- Kim, Nayeon. and Lee, Hyunsoo. (2021). Assessing consumer attention and arousal using eye tracking technology in virtual retail environment, *Frontiers in Psychology*, Volume 12.
- Krupas, Maros. Kajati, Erik. Liu, Chao. and Zolotova, Iveta. (2024). Towards a human-centric digital twin for human-machine collaboration: A review on enabling technologies and methods, *Sensors*, Volume 24, No. 7.
- Kumar, Rohit. and Dhar, Debayan. (2023). Unraveling the Potential of Immersive Virtual Environments for Behavior Mapping in the Built Environment: A Mapping Review, *Human Behavior and Emerging Technologies*, Volume 2023.
- Lanier, Madison. Waddell, T Franklin. Elson, Malte. Tamul, Daniel J. Ivory, James D. and Przybylski, Andrew. (2019). Virtual reality check: Statistical power, reported results, and the validity of research on the psychology of virtual reality and immersive environments, *Computers in Human Behavior*, Volume 100. pp. 70–78.
- Stefan, Hans. Mortimer, Michael. and Horan, Ben. (2023). Evaluating the effectiveness of virtual reality for safety-relevant training: A systematic review, *Virtual Reality*, Volume 27, No. 4. pp. 2839–2869.
- Zheleva, Aleksandra. De Bruyne, Jonas. Durnez, Wouter. Van De Walle, Sam. Verreyken, Siemon. Demanet, Jelle. and Bombeke, Klaas. (2023) “CaliBrainVR: Using Psycho-physiological Measures to Calibrate Virtual Reality Training”, proceedings of the 2023 15th International Conference on Quality of Multimedia Experience (QoMEX). pp. 123–126.