International Journal of Serious Games

ISSN: 2384-8766
https://journal.seriousgamessociety.org/

Article

Safe Walk: A Serious Game for Exploring Environmental Distractions Affecting Pedestrian Safety

Yoones Sekhavat[1,2], Joseph Mani[1], Seyed Vahid Mostafavi[2], Samad Roohi[2,3], Zhen Liu[4], Yazdan Movahedi[2]

[1]Deaprtment of Mathematics & Computer Science, Modern College of Business and Science, Muscat, Oman; [2]Faculty of Multimedia, Tabriz Islamic Art University, Tabriz, Iran; Computer Science & Information Technology, La Trobe University, Melbourne, Australia; Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, China
{yoones.sekhavat, drjosephmani} @mcbs.edu.om ; {sv.mostafavi, y.movahedi} @tabriziau.ac.ir; {s.roohi} @latrobe.edu.au; {liuzhen} @{nbu.edu.cn}

Keywords:

Serious Game
Pedestrian Safety
Environmental Distractions
Walking Simulator

Received: February 2025
Accepted: June 2025
Published: July 2025
DOI: 10.17083/02xf3365

Abstract

Understanding the impact of distracting factors on pedestrian safety is crucial for reducing accidents. This study investigates how urban video advertisements influence pedestrian behavior amidst competing visual stimuli. We employed a true experimental design, utilizing eye-tracking technology within a simulated environment to assess participants' visual behavior in the presence of video advertisements. Participants were randomly assigned to one of three scenarios, and we compared attention direction using both an eye-tracker and a virtual reality headset. The eye-tracking data captured from a two-dimensional screen was translated into three-dimensional coordinates for accurate representation in the VR environment. Our findings reveal a significant preference for monitoring vehicular traffic over video advertisements, confirming that participants prioritize safe navigation over engaging with distractions. This research advances the state of the art by demonstrating the reliability of eye-tracking in capturing attentional focus and highlighting the effectiveness of simulated environments for studying pedestrian behavior. The results have important implications for urban planning and safety measures, emphasizing the need to consider environmental distractions in the design of urban spaces.

1. Introduction

Pedestrian movement in urban environments is a vital component of social and economic interactions, facilitating connectivity and mobility within city life. As one of the most sustainable forms of transportation, pedestrian traffic represents a crucial element of urban mobility, second only to motor vehicles. A significant proportion of urban residents rely on walking as a mode of transportation at least once daily. To promote walking, urban infrastructure must address the social, psychological, and physiological needs of pedestrians

[1]. However, the increasing prevalence of motor vehicles, coupled with a neglect of pedestrian requirements in road design and land use, has rendered pedestrians particularly vulnerable to road accidents.

Among the various modes of transportation, pedestrians are the most susceptible to harm, as motor vehicle collisions frequently result in severe injuries or fatalities. According to traffic police reports, approximately 70% of fatalities in traffic accidents involve pedestrians [2]. This statistic underscores the urgent need for effective interventions to enhance pedestrian safety. Pedestrian traffic accidents arise from a complex interplay of factors, including road configuration, vehicle presence, and the interactions between road users. Consequently, ensuring pedestrian safety as the most vulnerable category of road users becomes a pressing concern.

High-risk locations for pedestrian accidents predominantly include intersections and crosswalks, where distractions from external factors and inattentiveness to traffic flow are the primary contributors to accidents [3]. External distractions may encompass the use of portable electronic devices, such as music players and mobile phones, as well as urban infrastructure elements, including traffic control systems, environmental advertisements, and construction sites [4]. These distractions not only divert attention from potential hazards but also contribute to a general lack of awareness of the surrounding environment, compounding the risks faced by pedestrians.

Among the various environmental distractions that can affect pedestrian safety such as smartphone use, social interactions, and auditory stimuli, this study specifically focuses on video advertisement screens placed in urban settings. These visual stimuli are increasingly common in modern cities and are strategically positioned in areas of high foot traffic, making them particularly relevant for examining their potential impact on pedestrian attention. Unlike personal distractions such as mobile phone use, which are self-initiated, video advertisements are externally imposed and often designed to capture attention through movement, brightness, and dynamic content. By isolating this specific type of distraction, the study aims to provide a clearer understanding of how urban design elements influence pedestrian behavior, offering valuable insights for city planners and public safety initiatives.

This research aims to investigate the visual behavior of individuals in traffic scenarios, utilizing eye-tracking technology as a methodological approach. Eye-tracking offers researchers valuable insights into visual attention by quantifying various eye movement parameters, such as gaze direction and fixation duration [5]. This technology is applicable across diverse populations, including infants, individuals with neurological disorders, and even animals [6, 7]. The natural feedback provided by eye-tracking enables a nuanced understanding of visual behavior in real world contexts, allowing researchers to identify not only where individuals direct their attention but also how long they remain focused on specific elements within their environment.

Given the inherent risks associated with studying pedestrian behavior in actual traffic environments, one objective of this research is to evaluate the feasibility of employing a simulated virtual environment as a safe and reliable platform for such studies. By utilizing this virtual environment, it becomes feasible to extract data on participants' visual behavior and movement patterns. The simulated setting facilitates the observation and analysis of participant interactions within a controlled context. The second objective of this research is to identify specific environmental factors that divert attention, leveraging the capabilities of eye-tracking technology.

While comparing real-world pedestrian behavior with simulated data could provide valuable insights into how accurately the simulation reflects actual pedestrian responses to environmental distractions, there are several challenges associated with this approach. Collecting real-world data involves complexities such as ethical concerns, uncontrolled variables, and technological limitations (e.g., the difficulty of replicating the precise control over experimental variables that is possible in a simulated environment). These challenges may have influenced the decision to focus on a controlled, simulated setting in this study. Nevertheless, future research should aim to bridge this gap by incorporating real-world data to validate the findings of simulated studies and further explore pedestrian behavior in more dynamic and unpredictable urban contexts.

In light of the rising prevalence of video advertising in urban settings, a critical question emerges regarding the impact of such advertisements on pedestrian safety and their potential to distract pedestrians from surrounding traffic events. This constitutes the third objective of the current study. The pervasive nature of digital advertising raises important questions about its role in shaping pedestrian behavior and awareness, particularly in high traffic areas where competing visual stimuli can lead to increased accident risk.

To operationalize this research, consultations with experts from the Traffic Department of Tabriz Municipality's Transport and Traffic Organization informed the modeling and simulation of the Tabriz Municipality Square intersection as a cohesive virtual environment for testing. This environment incorporates dynamic elements such as moving vehicles, an active construction site, traffic control facilities, and video advertisements, thereby enhancing participants' sense of immersion. Three intelligent traffic scenarios were designed for participant engagement, allowing for the collection of comprehensive data regarding participants' positions, directions, and visual behaviors throughout the simulated experiences.

This research contributes to the understanding of pedestrian safety by examining the effects of environmental distractions within a simulated context, providing insights that may inform urban planning and traffic safety strategies. By identifying the specific factors that divert attention and analyzing their implications for pedestrian safety, this study aims to offer evidence-based recommendations for designing urban spaces that prioritize the needs of pedestrians. The findings of this research could ultimately lead to the development of targeted interventions and policies that enhance the safety and wellbeing of pedestrians in urban environments, fostering a more sustainable and pedestrian-friendly urban infrastructure.

2. Related work

In their study evaluating human behavior at pedestrian crossings, Mako concluded that infrastructure, protective systems, and education significantly contribute to reducing the number and severity of pedestrian accidents [3]. Pedestrian crossings equipped with flashing lights, safe islands, and guide lights have had a positive impact on drivers' traffic flow and, consequently, reduced pedestrian accidents.

On the other hand, virtual reality (VR) refers to a system consisting of a computer with high-speed processing power, computer-generated graphics, and animations displayed through gloves or controllers equipped with position sensors, with its visual output displayed by a head-mounted display (virtual reality headset) [8]. According to [9], virtual reality is an environment composed of computer-generated images that respond to human movements. This virtual environment is experienced through goggles with dual displays and input gloves. All definitions include four elements: a computer as a processor, a head mounted display as a display, a locator for finding head position, and control gloves or controllers as input devices.

In a study on pedestrian visual perception and collision-avoidance behavior [10], several experiments were conducted in controlled laboratory environments to investigate the information processing space of pedestrians, which are extensively used in pedestrian motion simulation models. The findings of these experiments indicate that pedestrians allocate more attention to the ground level to detect potential hazards rather than obstacles; most fixation points of participants had a conical surface instead of a semicircular surface; the attention to other pedestrians was not equal to the attention to static obstacles. The results of this research suggest that the information processing structure needs to be reconsidered, taking into account the orientation characteristics of pedestrian vision.

In a study comparing the visual behavior of pedestrians and cyclists [11], eye-tracking was conducted in real world environments on pedestrians and cyclists. Participants traversed a predetermined route in the city either on foot or by bicycle, while eye-tracking data collected information related to their visual behavior. Additionally, accelerometer data and video recordings of the participants' field of view were analyzed. Shoulder glances, assumed as a visual behavior for safe passage, were identified from the recorded data. The analysis revealed that cyclists exhibited fewer shoulder glances and a shorter average duration compared to pedestrians, except during crossings. This shorter duration among cyclists was attributed to their higher speed, necessitating longer distances without forward-looking. Moreover, participants primarily focused on the route, with advertisements along the way having minimal distraction, and pedestrians garnering more attention than cyclists.

In a study on cyclists' visual behavior in urban environments [12], the visual behavior of cyclists on dedicated and shared paths with pedestrians was examined. Participants cycled predetermined routes while wearing eye-tracking devices. An algorithm was employed to detect fixation points and analyze frame-by-frame eye-tracking data. The correlation between fixation points and duration served as a measure of visual workload. The findings illustrated that physical and visual separation between pedestrians and cyclists enhanced cyclists' concentration, suggesting the potential benefits of segregated pathways.

Another study investigated eye movements during cycling in the game GTA V [13], utilizing the game's platform to study cyclists' traffic behavior. Participants engaged in various traffic scenarios within the game, with data on position, speed, and fixation points extracted and visualized in two separate maps. These maps provided insights into participants' attention allocation, with colors indicating speed variations and fixation points highlighting areas of focus. The study effectively showcased the utility of video games, coupled with eye-tracking, for researching human factors in traffic behavior.

In a study on crossing streets with a computer-generated virtual character in a virtual environment [14], researchers examined decision-making coordination against a virtual character in a street crossing scenario. This experiment took place in a virtual environment resembling a CAVE, featuring three vertically aligned display screens projecting the virtual environment stereoscopically using projectors. An immersive sound system replicated accurate traffic sounds. Results indicated a preference among participants to cross the street with a companion across all scenarios. Specifically, 64% of crossings involved a safe companion, 72% with a risky companion, and 76% with a real companion. Interestingly, participants tended to cross when the distance between vehicles was greater. Notably, no significant difference was observed in the distance between participants and real or virtual companions. In all scenarios, participants decided to cross the street simultaneously with their companions, entering the street together.

In [15], gamification is leveraged to enhance environmental awareness, showing how audio cues can improve navigation skills in real-world settings through an interactive audio game. Building on this theme, [16] presents a hierarchical architecture for planning vehicle behavior in serious games, employing deep reinforcement learning to enhance decision-making efficiency and safety while accommodating diverse driving styles. Similarly, [17] demonstrates that a cognitive load and performance-based system can boost player engagement and learning in environmental contexts by implementing real-time game modifications. Together, these studies illustrate the potential of gamification to enrich environmental interactions across different gaming scenarios.

Given the significance of pedestrian safety and the inherent risks involved in field studies, this research was conducted in a simulated virtual environment. Tabriz Municipality Square intersection was accurately modeled and simulated as a cohesive virtual environment for experimentation. Various elements such as different types of vehicles moving in city streets, a construction site, traffic facilities, video advertisements, and other immersive elements were integrated into this virtual environment. Three intelligent traffic scenarios were designed for testing, and participants actively engaged in these scenarios. Data regarding position, direction, and visual behavior of the participants were meticulously extracted and utilized in the final analysis.

Eye gaze tracking holds significant potential in cognitive behavioral studies. Its applications extend to human-computer interaction as an input device and encompass research in psychology, linguistics, health, and marketing. Furthermore, eye-tracking finds relevance in various social neuroscience issues [18], consumer marketing research [19], and user experience [20]. Research using eye-tracking in healthy individuals has unveiled tendencies such as focusing on eyes when confronted with social stimuli and perceiving faces holistically. Additionally, it has shed light on attention patterns in preverbal children, demonstrating a natural inclination towards eye fixation, except when speech learning occurs, diverting focus to the mouth area [21]. A recent study using eye tracking in a two-player racing game showed that real-time awareness of an opponent's emotions—conveyed through visual cues—can influence a player's emotional state and gameplay behavior, highlighting the role of affective interaction in online games [22]. A game-based study using gaze tracking showed that environmental distractions had limited impact, as pedestrians focused more on vehicles and the road than on advertising displays [23].

3. Safe Walk: Proposed Pedestrian Simulation Game

To investigate participant behavior, a pedestrian behavior simulator was developed. Participants engaged in three sequential scenarios within the simulator, allowing for the extraction of their behaviors.

In the first test scenario (Figure 1), participants navigate a one-way street where traffic flows from left to right. Here, participants must discern opportune moments to cross, considering the positions and distances of moving vehicles.

Figure 1. First scenario: navigating a one-way street where traffic flows from left to right (left), Starting point, endpoint, movement path in the test, and location of video advertisements (right).

In the second test scenario, the participant crosses a street where the initial section is a bus lane, followed by a section designated for vehicles. Here, a bus parked beside the pedestrian crossing obstructs the participant's view (Figure 2). As the participant enters the pedestrian crossing, another bus begins to move from beside the stationary bus (Figure 3). In this scenario, the participant must halt to prevent a collision and subsequently proceed to cross the bus lane once the bus clears the area. The latter part of this scenario mirrors the first scenario.

In the third test scenario, participants traverse a one-way street with a open manhole along the pedestrian crossing path (Figure 4). This section experiences heavy car traffic, adding complexity to the crossing. Participants must navigate safely, maintaining a suitable distance from passing vehicles, and avoid falling into the open manhole while crossing. Upon completing this street crossing, participants reach the endpoint of the test.

Figure 2. Second scenario: a bus obstructing the participant's view

Figure 3. Second scenario: a bus moving from beside the stopped bus

The simulated test environment was meticulously crafted in three dimensions. Initially, real world documentation, including photos and videos of the location, served as references for modeling the required 3D volumes in Maya software. Textures were then applied to these models, with detailed texturing carried out using Substance Painter, adhering to the Physically Based Rendering (PBR) system for realism.

Following the modeling and texturing phase, development transitioned to the programming and scenario design stage within the Unity game engine. Integration of the Tobii 4C eye-tracker into the Unity environment commenced with the installation of its development kit. Leveraging the libraries provided by this kit, a seamless connection between the eye-tracker and the game engine was established. Two-dimensional screen data captured by the eye-tracker was translated into three-dimensional coordinates within the Unity global coordinate system for accurate representation in the simulator's 3D environment.

Upon importing the meticulously crafted models and textures into the game engine, they were carefully placed within the simulator space. To facilitate analysis and reporting, significant objects within the environment were appropriately labeled. The control system of the simulator was implemented using a first person controller, providing participants with an immersive experience. Scenarios were crafted by strategically placing trigger volumes and collision planes in desired locations, ensuring realistic interactions. Additionally, car traffic within the environment was managed along specified routes, with distinct sound effects assigned to each vehicle. To enhance immersion, urban sound effects were carefully selected and integrated into the simulator environment.

Figure 4. Third scenario, an open manhole in the pedestrian crossing path

3.1 Identification of Dangerous Situations

To determine whether participants encountered dangerous or collision situations during the test, activating volumes were placed in various locations within the environment, particularly around vehicles. The risk of collision is influenced by the speed of the approaching vehicles. According to traffic regulations, the maximum speed limit for vehicles in roundabouts is 30 kilometers per hour; thus, the moving vehicles in the simulator operate at this speed.

Considering the speed of the vehicles, a collision zone is defined as two meters in front of the vehicle, with a danger zone extending two meters beyond that. Due to the curved path of vehicles in the roundabout, a dangerous area is also established 60 centimeters on either side of the vehicle (Figure 5).

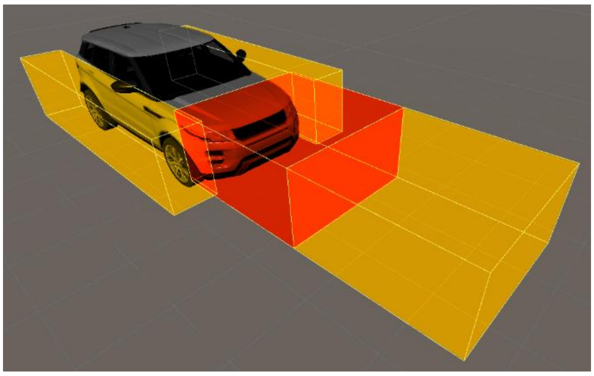

Figure 5. Placement of activating volumes to report danger positions (yellow) and collision positions (red) in front of and around the vehicles.

4. Evaluation

The independent variable in the test was the presence of video advertisements in the environment, a condition applied solely to the experimental groups. In the first scenario, participants' visual behavior was captured using an eye-tracker mounted below the display screen. Prior to data collection, the eye-tracker was calibrated for each participant, ensuring accuracy. Two-dimensional screen coordinates were received, which were then mapped onto the three-dimensional simulated environment using ray mapping functions. The system recorded the three-dimensional coordinates of gaze points, along with the label of the observed object, viewing time in seconds, and instances of being in a dangerous or collision situation.

In the second test, the participant's attention direction was directly mapped in the three-dimensional environment using the head tracking system of the virtual reality headset. This system recorded the three-dimensional coordinates of the attention points and the corresponding object labels in the report. Participants began the test by being placed in three predefined scenarios, and the test concluded after completion of the final scenario. Movement behavior of the participants was documented as three-dimensional coordinates in the report, with instances of being in a dangerous or collision situation also recorded.

Considering the objective of this research, which is to analyze pedestrian safety by simulating environmental distraction factors using eye-tracking technology, the test method employed is a true experimental design. Participants were randomly selected, and the design utilized is of the post-test type, featuring a control group alongside two independent experimental groups.

This study involving human participants was approved by the Research Ethics Committee of Tabriz Art University. Prior to participation, written informed consent was obtained from all participants. No identifying personal information has been included in the publication.

4.1 Participants

Participants in this study were recruited from the student body of the Faculty of Multimedia at Tabriz Islamic Art University, which comprised approximately 250 students across various majors at the time of the study. Convenience sampling was employed to select participants, with all students eligible for inclusion. A total of 40 undergraduate students were recruited, divided into four groups of 10 individuals each. The average age of participants was 28 years (ranging from 20 to 35), with a gender distribution of 42.5% female and 57.5% male across both test sessions. In terms of academic background, 50% were undergraduate students in computer arts and 50% were from multimedia programs. The participants had a relatively high average visual acuity of 2.85 out of 3, as measured by a standardized visual assessment used in the study. Notably, 65% of participants reported prior experience with virtual reality technologies. This factor is particularly relevant, as prior exposure to VR can influence levels of immersion, comfort, and task performance during virtual simulations. The demographic and experiential diversity of the participant pool contributes to the robustness of the findings and reflects a reasonable representation of the wider student population engaged in digital and immersive media studies. Overall, the selection of participants aimed to provide a representative sample from the Multimedia Faculty, ensuring diversity across the experimental and control groups for comprehensive analysis.

In the experimental groups, participants were exposed to the independent variable, namely video advertisements. One experimental group underwent testing in front of a screen equipped with an eye-tracker, while the other conducted the same test in a virtual reality environment. Conversely, the control groups did not experience the independent variable. Similar to the experimental groups, one control group completed the test in front of a screen with an eye-tracker, while the other conducted the test in a virtual reality setting.

For the first test, 10 participants from the experimental group (E1) and 10 from the control group (C1) were selected. This test involved conducting tasks in front of a screen equipped with an eye-tracker. Similarly, for the second test, 10 participants from the experimental group (E2) and 10 from the control group (C2) were chosen. This test took place in a virtual reality environment.

4.2 Data gathering and analysis tools

The Tobii Eye-Tracker 4C was used to extract visual behavior data from participants in the first test group, while the HTC Vive headset was employed for the second test group to present the virtual reality environment. The eye-tracker captures the participant's gaze point as two-dimensional coordinates on the screen. It is important to note that the VR scenarios did not include integrated eye-tracking within the virtual environment itself. Instead, the eye-tracking data was captured from a two-dimensional screen and then mapped into three-dimensional coordinates to accurately represent participants' visual focus within the VR simulation. This approach allowed for precise tracking of attentional focus while preserving the immersive qualities of the virtual environment.

To analyze the raw data collected from both the experimental and control groups, SPSS software was employed. Initially, the normal distribution of the data was assessed using the Kolmogorov-Smirnov test. Since the significance level (sig) for all parameters exceeded 0.05, indicating normal distribution, Pearson correlation analysis was employed to investigate the correlation between variables. Additionally, an independent t-test was conducted to identify significant differences between the two groups.

4.3 Procedure

In the first test, participants sit at a computer system fitted with an eye-tracker positioned beneath the display, and the calibration process is initiated. Calibration is performed individually for each participant. Following calibration, participants enter a simulated environment and spend 5 minutes in a controlled space to acquaint themselves with the input mechanisms and movement controls. These input tools include a mouse for controlling head movement and a keyboard for navigating the virtual agent within the simulated environment. Upon completion of the familiarization period, participants proceed to the main test environment. Throughout the test, data on participants' position (in 3D coordinates), gaze point (in 3D coordinates), the object viewed, and any instances of danger or collision are recorded.

In the second test, calibration settings for the virtual reality environment are adjusted before participants enter. Participants wear a virtual reality headset and spend 5 minutes in a controlled space familiarizing themselves with the simulator control system. In this test, navigation within the environment is facilitated by a gamepad. After the familiarization period, participants enter the main test environment and interact with the simulator according to provided instructions. Throughout the test, data on participants' position (in 3D coordinates), point of attention (in 3D coordinates), the object at the point of attention, and any instances of danger or collision are recorded. It is noteworthy that two participants experienced dizziness and headaches due to the virtual reality environment and were replaced by two other participants.

Upon completion of the tests, all participants responded to questions from two standard questionnaires: the Immersion Questionnaire [24] and the Engagement Questionnaire [25]. Both questionnaires utilized a Likert scale. Additionally, participants were asked to provide responses to two open-ended questions concerning the identification of distracting factors and the role of video advertisements as a distraction factor. Following the completion of the questionnaires and open-ended inquiries, the test session concluded, and participants exited the testing venue.

5. Results

This section outlines the methodology, calibration process, data collection procedures, and preliminary findings, tailored for an academic audience. The information and raw data acquired from the research tools utilized on both the experimental and control groups underwent analysis using SPSS, MATLAB, and Excel software. The analysis of the obtained data encompassed both descriptive and inferential levels.

For the initial test, the Kolmogorov-Smirnov test was employed to assess data distribution. With a significance level (sig) exceeding 0.05, the data distribution was deemed normal (Table 1). Given the normal distribution, the inferential results section employed the Pearson correlation test to scrutinize variable correlations, while the independent t-test was utilized to identify significant differences between the two groups.

Table 1. Results of the Kolmogorov-Smirnov Test on Data Extracted from the First Test

Variable	Mean	Standard Deviation	Sig. Level (P)
Duration of Watching Video Advertisements	2.98	3.09	0.154
Time Spent in Risky Conditions (Test)	1.03	0.59	0.200
Time Spent in Accident Conditions (Test)	0.37	0.37	0.570
Time Spent in Risky Conditions (Control)	1.59	0.56	0.200
Time Spent in Accident Conditions (Control)	0.58	0.53	0.170

For the second test, data distribution was assessed using the Kolmogorov-Smirnov test. With a significance level (sig) surpassing 0.05, the data distribution was deemed normal (Table 2). Given the normality of the data, the inferential results section employed Pearson correlation analysis to examine variable correlations and the independent t-test to identify significant differences between the two groups.

Table 2. Results of the Kolmogorov-Smirnov test on the data extracted from the second test

Variable	Mean	Standard Deviation	Sig. Level (P)
Duration of Video Advertisements Viewed	1.66	1.48	0.20
Time Spent in Hazardous Conditions (Test)	1.6	0.86	0.25
Time Spent in Accident Conditions (Test)	1.32	1.7	0.15
Time Spent in Hazardous Conditions (Control)	1.49	0.78	0.36
Time Spent in Accident Conditions (Control)	1.22	0.76	0.41

After clustering the raw data of three-dimensional coordinate points of views based on temporal proximity in MATLAB software, a score is assigned to each cluster based on its membership count. This data is then paired with the corresponding three-dimensional coordinate positions of the clustered views. The three-dimensional coordinate data are mapped onto collision surfaces to extract two-dimensional coordinates of the spatial texture on these surfaces.

Subsequently, the processed data is outputted as a text file containing paired coordinates of views and positions, along with a heat map generated by MATLAB software. The Unity game engine is then employed to visualize the data. The heat map image is projected onto the screen, with more-viewed points depicted in yellow and less-viewed points shown in blue. Meanwhile, the text data undergo processing, reading, and display, showcasing the views with green volumes (Figure 6). The size of these volumes reflects the score of the corresponding point, indicating prolonged attention at that specific location. Additionally, blue lines are used to connect the green volumes to the points from which they were viewed.

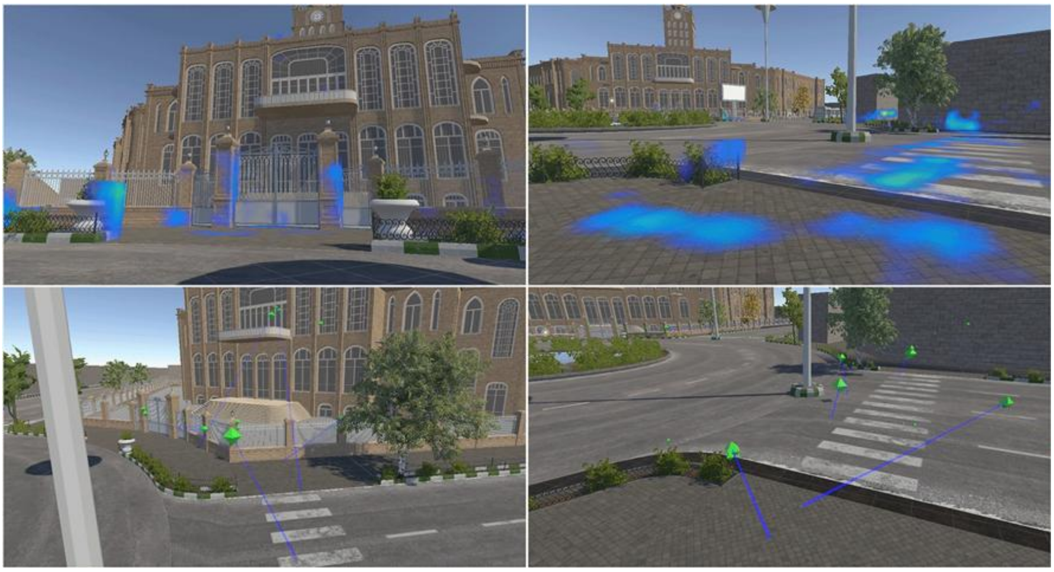

Figure 6. Mapped heat map on the three-dimensional environment (top), visualization of clustered viewing points of a participant (bottom).

After segregating the raw data and obtaining the duration of object views by each participant, box plots were generated for two factors in the first test: the duration of watching video advertisements and the duration of viewing cars (Figure 7). The plot on the right side (diagonal hatch) illustrates the distribution of data for the duration of viewing cars, while the plot on the left side (vertical hatch) represents the distribution of data for the duration of watching video advertisements.

Observing the box plots reveals that cars, crucial for safely crossing pedestrian lanes, were viewed significantly more, indicating a substantial disparity between the data of the two columns with no overlap. Thus, it can be inferred that the designed simulator offers a reliable platform for studying pedestrian traffic behavior.

The presented results demonstrate a significant difference between the average durations of watching video advertisements and viewing cars in the experimental group. These findings suggest that participants prioritize the safe crossing of pedestrian lanes, placing greater emphasis on cars.

Analysis of the raw data, which includes labeled viewed objects over time alongside the presence of video advertisement labels, highlights the eye-tracker's effectiveness in detecting distracting environmental factors. The eye-tracker proves to be an efficient tool for detecting and recording these factors, thereby confirming the second hypothesis.

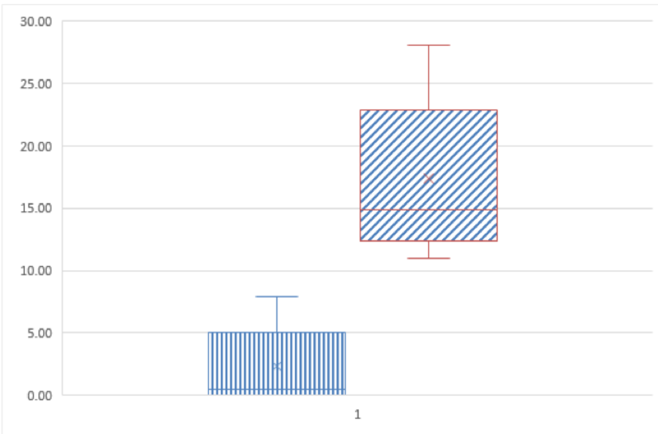

Figure 7. Box plots for two factors, the duration of watching video advertisements and the duration of viewing cars.

In the first test, Pearson correlation analysis was employed to examine the correlation between variables, given the normal distribution of the data. The variables scrutinized were the duration of watching video advertisements, the duration of being in risky and accident conditions for each participant in the experimental group. The results of this analysis are presented in Table 3.

Table 3. Correlation between Time Spent in Hazardous and Accident Conditions

Variable	Time in Hazard Conditions	Time in Accident Conditions
Duration of Viewing Ads	0.178	0.324
Significance Level (P)	0.673	0.433

The results from the Pearson correlation analysis reveal a weak positive correlation coefficient of 0.178 between the duration of watching video advertisements and the duration of being in risky conditions. With a significance level (P) greater than 0.05, this indicates no significant relationship between these variables. Similarly, the correlation coefficient between the duration of watching video advertisements and the duration of being in accident conditions is moderate, at 0.324, with a significance level of 0.433, signifying no significant relationship between these variables.

To compare the durations of being in risky and accident conditions between the experimental and control groups in the first test, the independent t-test is employed, given the normal distribution of the data (Table 4).

Table 4. Results of independent t-test in the first test

Variable	Independent t (t)	Significance Level (P)
Duration of Exposure to Hazardous Situation	2.60	0.056
Duration of Exposure to Accident Situation	0.94	0.36

The results from the independent t-test indicate a significance level of 0.056 for the duration of being in risky conditions and 0.36 for the duration of being in accident conditions between the experimental and control groups. Since both values exceed 0.05, there is no significant relationship between these two groups.

In the second test, Pearson correlation analysis is employed to examine the correlation between variables, given the normal distribution of the data. The variables analyzed include the duration of watching video advertisements, the duration of being in risky conditions, and the duration of being in accident conditions for each participant in the experimental group. The results of this analysis are presented in Table 5.

Table 5. Results of Pearson correlation analysis in the second test

Variable	Time in Hazard Conditions	Time in Accident Conditions
Duration of Viewing Ads	-0.494	0.174
Significance Level (P)	0.319	0.742

The results from the Pearson correlation analysis reveal a moderate negative correlation coefficient of -0.494 between the duration of watching video advertisements and the duration of being in risky conditions. However, since the significance level (P) exceeds 0.05, there is no significant relationship between these two variables. Similarly, the correlation coefficient between the duration of watching video advertisements and the duration of being in accident conditions is weak, at 0.174, with a significance level of 0.742, indicating no significant relationship between these variables.

To compare the durations of being in risky and accident conditions between the experimental and control groups in the second test, the independent t-test is employed, as the data follows a normal distribution.

Table 6. Results of the independent t-test in the second test

Variable	Independent t (t)	Significance Level (P)
Duration of Exposure to Hazardous Situation	0.256	0.803
Duration of Exposure to Accident Situation	0.124	0.904

The results obtained from the independent t-test (Table 6) indicate significance levels of 0.803 for the variable of time spent in risky conditions and 0.904 for the variable of time spent in accident conditions in both the test and control groups. Since both of these values exceed 0.05, there is no significant relationship between these two groups. These findings suggest that the duration of watching video advertisements does not significantly affect the duration of exposure to risky and accident conditions ($P > 0.05$) in both tests. Additionally, there is no significant relationship between the duration of being in risky and accident conditions across the two test and control groups ($P > 0.05$). Therefore, the presence of video advertisements does not divert pedestrian attention from surrounding traffic events and does not increase the likelihood of accidents.

Overall, the results from both tests demonstrate no significant relationship between the duration of watching video advertisements and the time participants spent in either risky or accident conditions. Across all correlation analyses, the coefficients were weak to moderate but not statistically significant ($P > 0.05$), indicating no meaningful association. Similarly, the independent t-tests showed no significant differences between the experimental and control groups in terms of exposure to risky or accident conditions. These findings suggest that video advertisements did not significantly affect pedestrian behavior or increase the risk of accidents in the simulated environment.

To examine the normal distribution of the questionnaire data on immersion and engagement, the Kolmogorov-Smirnov test was employed. As shown in Table 7, the significance level for all variables is reported to be higher than 0.05, indicating that the data follows a normal distribution.

Table 7. Results of the Kolmogorov-Smirnov test on immersion and engagement questionnaire data

Variable	Mean	Standard Deviation	Significance Level (P)
Immersion (Test 1)	117.75	8.53	0.138
Engagement - Absorption (Test 1)	20.95	2.11	0.212
Flow (Test 1) Engagement	16.5	2.81	0.282
Presence (Test 1)	16.85	2.74	0.298
Immersion (Test 2)	126.5	8.23	0.461
Engagement - Absorption (Test 2)	21.58	3.17	0.337
Engagement - Flow (Test 2)	19.25	1.91	0.061
Engagement - Presence (Test 2)	18.92	1.78	0.232

By examining the mean data from the engagement questionnaire, it is observed that participants in the second test (virtual reality) exhibited higher engagement with the simulated virtual environment across all factors, including absorption, flow, and presence.

Table 8. Results of the independent t-test comparing questionnaire data between the first and second tests

Variable	Duration of Viewing Ads	Significance Level (P)
Immersion	2.84	0.008*
Engagement – Absorption	0.68	0.5
Engagement – Flow	2.98	0.006*
Engagement - Presence	2.32	0.027*

By examining the results presented in Table 8, it is evident that a statistically significant relationship exists between the variables of the test for all factors, with the exception of absorption (significance level = 0.05). Specifically, the p-values for these factors are less than 0.05, indicating a strong correlation between the tested variables. This analysis highlights the effectiveness of the testing conditions in capturing meaningful differences.

Moreover, based on the test results and the associated mean data, the second test utilizing virtual reality demonstrated superior performance in terms of immersion, flow, and presence when compared to the first test conducted with an eye-tracker positioned under the monitor. The enhanced immersive experience provided by the virtual reality setup likely contributed to participants feeling more engaged and present within the simulated environment. This finding suggests that the use of virtual reality can significantly enhance the quality of experiential research, offering deeper insights into user behavior and interactions.

6. Discussion

The results of this study confirm the initial hypothesis, showing that participants consistently prioritized monitoring vehicular traffic over video advertisements. This preference for focusing on immediate traffic hazards over visual distractions is in line with existing research on pedestrian behavior, which emphasizes the importance of attention to potential risks in the environment, particularly in high-risk locations such as pedestrian crossings [10]. The clear trend observed across all test groups further supports the notion that, despite the visually stimulating nature of video advertisements, pedestrians instinctively prioritize safe navigation, a finding that can inform urban planning strategies aimed at enhancing pedestrian safety.

The effectiveness of eye-tracking technology in capturing attentional focus within this simulated environment is evident, with the data showing high reliability. Previous studies utilizing similar technology, such as those examining eye movements during virtual cycling or in video game landscapes [13], support the accuracy and robustness of the eye-tracker in assessing participants' visual behavior in controlled settings. This underscores the utility of eye-tracking as a tool for studying pedestrian attention in simulated environments, offering an accurate portrayal of real-world attentional dynamics.

Furthermore, the absence of a significant correlation between video advertisement exposure and attention to risky conditions further reinforces the conclusion that pedestrians remain primarily focused on their surroundings rather than distractions. This finding is consistent with research highlighting the tendency of individuals to allocate their visual attention toward their movement trajectory and immediate environment, rather than external stimuli such as advertisements [11]. The overall implications of this are critical for urban safety initiatives, suggesting that while environmental distractions like video advertisements are ubiquitous in cities, pedestrians are generally more attuned to immediate safety concerns.

The immersive quality of the virtual reality (VR) environment used in this study also played a key role in enhancing participant engagement and sense of presence, as indicated by the questionnaire results. The VR platform offered a more interactive and engaging experience compared to traditional screen-based presentations, contributing to higher immersion levels. This heightened sense of presence not only fostered greater participant involvement but also enhanced the effectiveness of the study in simulating real-world pedestrian behavior. The results suggest that VR simulations can be a powerful tool for understanding and improving pedestrian safety, providing a safe yet effective means of investigating pedestrian behavior in high-risk urban scenarios.

While this study offers valuable insights into the impact of urban video advertisements on pedestrian behavior within a simulated environment, several limitations should be acknowledged to contextualize the findings and guide future research efforts. First, the relatively small sample size may have limited the statistical power of the analyses and reduced the generalizability of the results. Although the participant group provided valuable data on visual attention in a controlled experimental setting, expanding the sample to include a more diverse population would allow for more robust conclusions. Future studies should include participants from varying age groups, educational backgrounds, and levels of familiarity with virtual environments to better reflect the heterogeneous nature of urban pedestrians. Second, the use of a virtual reality-based environment, while beneficial for controlling variables and ensuring participant safety, presents certain limitations in ecological validity. The simulated environment lacked some real-world complexities such as ambient sounds, weather conditions, and the unpredictable behavior of other pedestrians or vehicles. These contextual elements can significantly influence pedestrian attention and decision-making. Therefore, future work should consider integrating more realistic environmental features or complementing VR studies with field-based experiments to validate findings under real-world conditions. Third, although the eye-tracking system within the VR headset effectively captured attentional focus, technical limitations such as calibration drift, tracking delays, and participant discomfort may have introduced minor inconsistencies in the data. These issues are particularly relevant in longer sessions, where user fatigue and hardware sensitivity can affect measurement accuracy. Future implementations should explore improved eye-tracking hardware, better calibration protocols, and more comfortable VR setups to enhance data reliability.

Additionally, while the general trend observed in this study—participants prioritizing monitoring traffic over attending to video advertisements—was consistent, not all condition-specific comparisons reached statistical significance. This suggests that attentional behavior in complex visual environments may vary in subtle ways depending on contextual or individual factors. Further research is needed to explore these nuances, potentially through adaptive or personalized experimental designs that account for individual traits such as risk tolerance or visual sensitivity. Finally, the study focused solely on short-term behavioral responses to visual distractions. However, pedestrian attention and behavior may evolve over time due to habituation, changing perceptions of risk, or increased familiarity with the environment. Longitudinal studies are therefore necessary to assess whether repeated exposure to video advertisements leads to desensitization, heightened caution, or other adaptive behaviors. For example, while pedestrians may initially disregard advertisements in favor of traffic monitoring, over time they may shift their attention allocation in more complex ways. Long-term observations could also reveal changes in risk compensation behaviors or engagement levels that are not detectable in single-session experiments. Incorporating repeated exposures and extended observation periods in future studies would provide a more comprehensive understanding of how urban distractions shape pedestrian safety over time. Addressing these limitations in future research will be crucial for enhancing the realism and generalizability of VR-based pedestrian studies and for informing more effective urban design and public safety policies.

Moving forward, it is recommended that future studies explore innovative methods for integrating video advertisements and other urban elements into simulations. Utilizing VR headsets equipped with eye-tracking capabilities and organic motion systems could enhance immersion while minimizing technical barriers for participants, ultimately providing a more authentic and immersive experimental experience. Additionally, efforts to visualize data from similar studies can offer valuable insights for optimizing the integration of urban elements in simulated environments.

7. Conclusions

This study highlights the critical interplay between pedestrian behavior and environmental distractions, particularly in urban settings where video advertisements are prevalent. The findings confirm that pedestrians prioritize safety over engaging with advertising content, as evidenced by their attention being predominantly directed towards navigating their immediate surroundings, a pattern consistent with existing literature on collision avoidance and attentional focus on ground-level hazards [10, 11]. Notably, the lack of correlation between attention to advertisements and exposure to high-risk scenarios further supports the notion that pedestrians actively maintain situational awareness to mitigate potential threats.

The research also demonstrates the efficacy of eye-tracking technology in capturing nuanced aspects of visual behavior within simulated traffic environments. The eye-tracker's reliability aligns with its proven utility in similar studies, such as those investigating eye movements in virtual cycling and gaming contexts [13], reinforcing its applicability in urban pedestrian research. Participants' high levels of engagement and presence within the VR setting suggest that immersive simulations provide a valuable and safe platform for analyzing pedestrian interactions in realistic scenarios.

However, the study also acknowledges limitations, including potential motion sickness and the lack of integrated eye-tracking systems in some VR headsets. Addressing these issues in future studies, such as through the use of organic motion systems and advanced HMDs with built-in eye-tracking, could enhance immersion and data accuracy, reducing participant discomfort and increasing ecological validity. Additionally, visualizing data across similar experimental contexts may yield actionable insights for optimizing the integration of urban elements, including advertisements, within virtual environments.

Ultimately, this study provides valuable insights for urban planners and policymakers aiming to improve pedestrian safety. By identifying the factors that capture pedestrian attention and analyzing their implications, the research offers evidence-based recommendations for designing urban environments that prioritize pedestrian needs. Future investigations should aim to refine simulation techniques and explore innovative ways to integrate urban elements, thereby contributing to the development of safer, more pedestrian-friendly cities.

Acknowledgments

This research was supported by the Driving Simulator Lab at the Faculty of Multimedia, Tabriz Islamic Art University. We would like to thank all the anonymous participants who took part in the user study. Special thanks to Hesam Sakian Mohammdi for his valuable assistance in the development of the simulator.

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this study.

References

[1] M. M. Hamed, "Analysis of pedestrians' behavior at pedestrian crossings," *Safety Science*, vol. 38, no. 1, pp. 63–82, 2001. doi: 10.1016/S0925-7535(00)00058-8.

[2] National Research Council, *HCM 2010: Highway Capacity Manual*, Transportation Research Board, 2010. [Online]. Available: https://books.google.co.uk/books?id=rMofuQEACAAJ

[3] E. Mako, "Evaluation of human behaviour at pedestrian crossings," in *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 443–447, 2015. doi: 10.1016/j.trpro.2016.05.227.

[4] R. Rastogi, S. Chandra, and M. Mohan, "Development of level of service criteria for pedestrians," *Journal of the Indian Roads Congress*, vol. 75, no. 1, pp. 61–70, 2014. Available: http://hdl.handle.net/123456789/7690

[5] S. Almeida, O. Mealha, and A. Veloso, "Video game scenery analysis with eye tracking," *Entertainment Computing*, vol. 14, pp. 1–13, 2016. doi: 10.1016/j.entcom.2015.12.001.

[6] C. E. Venker and S. T. Kover, "An open conversation on using eye-gaze methods in studies of neurodevelopmental disorders," *Journal of Speech, Language, and Hearing Research*, vol. 58, no. 6, pp. 1719–1732, 2015. doi: 10.1044/2015_JSLHR-L-14-0304.

[7] C. J. Machado and E. E. Nelson, "Eye-tracking with nonhuman primates is now more accessible than ever before," *American Journal of Primatology*, vol. 73, no. 6, pp. 562–569, 2011. doi: 10.1002/ajp.20928.

[8] J. Steuer, "Defining virtual reality: Dimensions determining telepresence," *Journal of Communication*, vol. 42, no. 4, pp. 73–93, 1992. doi: 10.1111/j.1460-2466.1992.tb00812.x.

[9] R. Fernandes, A. P. Gaonkar, P. J. Shenoy, A. P. Rodrigues, B. A. Mohan & V. Padmanabha, (2021). Efficient virtual reality-based platform for virtual concerts. In Multimedia and Sensory Input for Augmented, Mixed, and Virtual Reality. doi: 10.4018/978-1-7998-4703-8.ch008.

[10] K. Kitazawa and T. Fujiyama, "Pedestrian vision and collision avoidance behavior: Investigation of

the information process space of pedestrians using an eye tracker," in *Pedestrian and Evacuation Dynamics 2008*, pp. 95–108, Springer, 2010. doi: 10.1007/978-3-642-04504-2_7.
[11] M. Trefzger, T. Blascheck, M. Raschke, S. Hausmann, and T. Schlegel, "A visual comparison of gaze behavior from pedestrians and cyclists," *Symposium on Eye Tracking Research and Applications*, 2018. doi: 10.1145/3204493.3204553.
[12] A. Mantuano, S. Bernardi, and F. Rupi, "Cyclist gaze behavior in urban space: An eye-tracking experiment on the bicycle network of Bologna," *Case Studies on Transport Policy*, vol. 5, no. 2, pp. 408–416, 2017. doi: 10.1016/j.cstp.2016.06.001.
[13] P. Bazilinskyy, N. Heisterkamp, P. Luik, S. Klevering, A. Haddou, M. Zult, and J. de Winter, "Eye movements while cycling in GTA V," 2018. Available: https://pure.tudelft.nl/ws/portalfiles/portal/68971745/EyemovementswhilecyclinginGTAV.pdf
[14] Y. Jiang, E. O'Neal, P. Rahimian, J. P. Yon, J. M. Plumert, and J. Kearney, "Action coordination with agents: crossing roads with a computer-generated character in a virtual environment," *Proceedings of the ACM Symposium on Applied Perception*, pp. 57–64, 2016. doi: 10.1145/2931002.293100.
[15] S. Poudratchi, Y. A. Sekhavat, M. R, Azadehfar, & S. Roohi. BatSight: A Navigation Game to Map Environmental Information into Audio Cues. *International Journal of Serious Games*, *11*(1), 101-119, 2024. doi: 10.17083/ijsg.v11i1.718.
[16] L. Forneris, A. Pighetti, L. Lazzaroni, F. Bellotti, A. Capello, M. Cossu, & R. Berta. *Implementing Deep Reinforcement Learning (DRL)-based Driving Styles for Non-Player Vehicles. International Journal of Serious Games. 10, 153–170 (2023)*. doi: 10.17083/ijsg.v10i4.638.
[17] A. Seyderhelm, & K. Blackmore. Dynamic Adaptive Surveillance Training in a Virtual Environment Using Real-Time Cognitive Load and Performance. *International Journal of Serious Games*, *11*(3), 109-133, 2024. doi: 10.17083/ijsg.v11i3.733.
[18] E. Birmingham and A. Kingstone, "Human social attention," *Annals of the New York Academy of Sciences*, vol. 1156, no. 1, pp. 118–140, 2009. doi: 10.1016/S0079-6123(09)17618-5.
[19] H. Khachatryan, A. Rihn, B. Campbell, C. Yue, C. Hall, and B. Behe, "Visual attention to eco-labels predicts consumer preferences for pollinator friendly plants," *Sustainability*, vol. 9, no. 10, p. 1743, 2017. doi: 10.3390/su9101743
[20] F. Ioannidou, F. Hermens, and T. L. Hodgson, "Mind your step: the effects of mobile phone use on gaze behavior in stair climbing," *Journal of Technology in Behavioral Science*, vol. 2, no. 3–4, pp. 109–120, 2017. doi: 10.1007/s41347-017-0022-6.
[21] E. Di Giorgio, D. Meary, O. Pascalis, and F. Simion, "The face perception system becomes species-specific at 3 months: An eye-tracking study," *International Journal of Behavioral Development*, vol. 37, no. 2, pp. 95–99, 2013. doi: 10.1177/01650254124653.
[22] Y. A. Sekhavat, S. Roohi, H. S. Mohammadi, and G. N. Yannakakis, "Play with one's feelings: A study on emotion awareness for player experience," *IEEE Transactions on Games*, vol. 14, no. 1, pp. 3–12, 2020. doi: 10.1109/TG.2020.3003324

[23] S. V. Mostafavi, Y. A. Sekhavat, S. Roohi, H. S. Mohammadi, and K. Pouralvar, "A game-based system to study the danger of advertising displays for pedestrians: Are they really dangerous?," in Proc. *2019 Int. Serious Games Symp*. (ISGS), Dec. 2019, pp. 68–73. doi: 10.1109/ISGS49501.2019.9047025
[24] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton, "Measuring and defining the experience of immersion in games," *International Journal of Human-Computer Studies*, vol. 66, no. 9, pp. 641–661, 2008. doi: 10.1016/j.ijhcs.2008.04.004.
[25] J. H. Brockmyer, C. M. Fox, K. A. Curtiss, E. McBroom, K. M. Burkhart, and J. N. Pidruzny, "The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing," *Journal of Experimental Social Psychology*, vol. 45, no. 4, pp. 624–634, 2009. doi: 10.1016/j.jesp.2009.02.016.