Human Interaction and Emerging Technologies (IHIET 2023), Vol. 111, 2023, 314–323
https://doi.org/10.54941/ahfe1004038

User Performance Analysis in Guided and Non-Guided Stressful Virtual Reality Scenarios

Alejandro Villar and Carlos León

Computer Science Faculty, Universidad Complutense de Madrid, 28040 Madrid, Spain

ABSTRACT

Virtual reality (VR) is a potentially useful technology for training and simulating hazardous scenarios. It is particularly useful for firefighters, healthcare staff or the military, which face emergency situations that involve high levels of stress. Studies have shown that the participant's stress level is increased when confronted with a hazardous scenario in a virtual environment, as would be expected in a realistic simulation. This research describes a methodology based on measuring interactions and objectives in hazardous scenarios, and an empirical study for assessing the participant's performance improvement. A prototype virtual reality experience in which the participants face a radiation emergency has been developed and a pilot study has been run. The methodology, a qualitative and a quantitative analysis with preliminary but promising results are described.

Keywords: Virtual reality, Virtual environment, Stress, Learning process, Detection, Video game, Emergency, Human computer interaction

INTRODUCTION

Multiple work environments use simulations to enhance learning. These simulations usually take place in environments where the user performs a series of controlled actions that are observed in order to assess the learning process. Simulations are particularly useful in learning for hazardous scenarios, ideally replicating stressful situations that might arise in them. Analysing the acquisition of skill in these potentially stressful environments is of importance to understand the user's capabilities and make progressive improvements.

Virtual Reality (VR) is used for entertainment, experimentation, and training in situations that could not be recreated in real life (Reznek, Harter and Krummel, 2002; Ruesseler *et al.*, 2010; Kaplan *et al.*, 2021). Virtual reality, as a particular tool for simulation, provides users with experiences that are like the non-simulated environment. Moreover, virtual reality is becoming more accessible to the general public and many teams are using it in the workplace or for training in emergency situations.

In a learning context, in addition to the relevance of the specific objectives, interactions play a fundamental role in VR. Interaction is an important factor for enhancing the realism of the virtual reality experience. To achieve useful and realistic interaction, the participant must be able to interact with the elements approximately in the same way as she does in reality. This study seeks to propose and test a methodology for learning assessment in VR hazardous environments. This methodology is based on the interactions with elements of the environment and fulfilment of objectives.

In particular, the methodology has been tested against a radiation emergency scenario. Assessing performance is usually a complex task, but radiation emergencies have a very clear set of protocols and there are several interactions that must be carried out a number of times, while others must not. Against this context, the current research reports on a methodology adapted to the training in radiation emergencies, and therefore is fully focused on the analysis of *interactions* and *objectives*. The present study consists of a VR experience in which participants were confronted with a radiation emergency. The participant's interactions and fulfilled objectives were annotated and compared with personal data obtained with questionnaires. These data were used to assess whether the participants had learned how to solve the proposed scenario.

PREVIOUS WORK

Virtual environments allow one to have control of the actions performed by the student at each moment. Using the gathered data, it is possible to carry out studies leading the improving to the learning process. Virtual reality is most often used to enable users to play video games where human-computer interaction is very close to reality, but the current state of the technology has led to a boom in studies that seek to demonstrate the relationship between video games and emotion change (Subahni, Xia and Malik, 2012; Porter and Goolkasian, 2019; Pine *et al.*, 2020). In addition, due to the pandemic caused by Covid-19, the use of this technology for people's entertainment has increased and therefore its relationship has also been studied (Pallavicini, Pepe and Mantovani, 2022). Studies show that stress influences people's decision making (Porcelli and Delgado, 2017; Wemm and Wulfert, 2017). Since stress tends to appear in risky situations, these decisions can have serious consequences (Starcke *et al.*, 2008). To avoid this, it is important to understand the person's behaviour at any given moment. In particular, risk situations such as catastrophes or radiological emergencies cause stress (Lopez Vazquez, 2001). For this reason, people who must face these emergencies must be trained beforehand. Stress directly affects the learning and memorization process (Joëls *et al.*, 2006; Schwabe *et al.*, 2012). Low and controlled levels of stress can help in this task, but high levels have the opposite effect.

Apart from entertainment, this technology can also help in learning processes (Freina and Ott, 2015; Kavanagh *et al.*, 2017), and in particular, groups of workers can be trained in more realistic simulations. It is therefore important to detect whether emotions in a virtual environment are the same in real-world situations.

Emergency situations are one of the most stressful situations for humans. This is because the person must face an unexpected situation that is possibly beyond her capabilities. It is important to detect these stress levels in order to understand human behaviour in these environments (Pluntke *et al.*, 2019; Ishaque *et al.*, 2020; Mevlevioğlu, Tabirca and Murphy, 2022). By doing so, actions that lead to a worse problem can be avoided. Several studies have sought to use virtual reality to improve people's health by managing stress and anxiety (Smith, Conway and Karsh, 1999; North and North, 2016; Soyka *et al.*, 2016; Emmelkamp, Meyerbröker and Morina, 2020). It has also been shown to overcome fears (Rothbaum *et al.*, 2000; Freeman *et al.*, 2018) and phobias (Botella *et al.*, 2017). To avoid these problems, simulations are created where people can face these situations without any danger (Keitel *et al.*, 2011). (Benoit *et al.*, 2009; Bouchard *et al.*, 2012; Klingner *et al.*, 2020) use this technology to detect the stress of their professionals in order to understand their actions. On the other hand, it allows the participant to learn new concepts (Checa, Miguel-Alonso and Bustillo, 2021) as it happens with health staff (Mantovani *et al.*, 2003), industrial mining (Van Wyk and De Villiers, 2009), or orthopaedic operations (Aïm *et al.*, 2016).

OBJECTIVES AND INTERACTIONS

As previously mentioned, the present research focuses on developing a methodology for assessing performance increase in a VR simulation of a hazardous scenario. In particular, the performance analysis is based on the analysis of *specific interactions* and the fulfilment of *concrete objectives* in the virtual environment.

The analysis is focused on checking whether an increase in performance between guided and non-guided stages of the same scenario can be evidenced. The experimental methodology includes a set of specific objectives that correspond to the protocol decisions that the participant must take in the virtual scenario. In addition to the completion of these objectives, the quality of each completion is measured through the number of interactions. This measurement is important in radiation emergency scenarios because the number of interactions defines the amount of time the participant has been potentially exposed to radiation and to manipulation of potentially radioactive material, which must be kept to a minimum.

Given the impact of specific details on the assessment of performance, the methodological analysis is fundamentally qualitative, based on a detailed analysis of the videos and the collected data, and objectives and interactions are studied separately to check if there are any variations between the qualitative and the quantitative findings. However, quantitative data gathered during the executions of the experience is also provided and analysed.

EXPERIMENTAL SETUP

An HTC Vive Pro 2 virtual reality wired headset was used for the experiment. Within the environment, the user was able to move and interact with the objects in the scenario. Movement was carried out by a teleportation mechanism in which the participant launches a beam that indicates the area to which she will be teleported, pressing a button on the controller. Realistic actions were used to interact with objects and characters. However, more complex actions, such as initiating conversations or opening doors, are performed by pointing at a beam of light.

The experience was designed for people not trained in radiation emergencies or the protocols to be followed. The aim is to analyse the learning process and results of participants who are unaware of the involved challenges. The participant was confronted with a traffic accident, vehicles driving along the road, characters to talk to in order to extract information, traffic cones, a radiation detector, boxes storing radioactive material, and a communication device to explain what has happened to the headquarters. The specific objectives were gathering information from characters, protecting and talking to the injured person, cutting off traffic, inspecting the van, checking the radiation levels and informing the head office.

These objectives had to be completed in order. However, sometimes they could be interspersed or even avoided. For example, it was important to ask witnesses about what happened, but it was not necessary to do so in a precise order. Sometimes actions like this were unnecessary, either because there were no witnesses or because sufficient information had been obtained from those involved. Another example was cutting off traffic using traffic cones. It was valid to cut off traffic before starting the inspection.

The possible interactions in the environment depended on the specific object of interaction. There were objects that could be picked up, such as traffic cones, characters could be talked to, etcetera. Interactions included: looking or stop looking, aiming or stop aiming, interacting with an object, using correctly or incorrectly an object, picking up or dropping an object, positioning correctly or incorrectly, talking to a character, and having a conversation with a virtual character.

The experimental process was divided into two stages for comparison. In the first stage, the participant was guided by means of information panels that explained the steps to follow. In the second stage, the participant had to complete all the objectives, demonstrating that the skill introduced in the first part had been acquired, without any guidance. The second stage had the same elements as the first stage, and only the location and the layout of elements had changed. The participant had to try to replicate all the steps in the same order to be successful in the experiment.

At the beginning of the experiment, the participants had 15 minutes where they were informed about the objective of the study and informed that they could end their participation at any time. Once informed, the participants had to fill out a questionnaire gathering personal data, which included information about the participants' skills or experience in virtual reality systems. Once the questionnaire had been completed, the participants were informed about the virtual setup, and they were briefed about the possibility of a simulated radiation leakage. The participants were also informed about the potential risks for the virtual characters. Before starting the experience, the controls were explained to the participant. After having learnt how to move and interact in the environment, the participants had 3 minutes to test the interface. After that, the participants started the guided stage. The participant had to describe her subjective experience and had a break of 1 minute and 30 seconds. After this time the participant started the non-guided part.

This phase did not have an explicitly set duration, as each player would need a different amount of time to finish.

ANALYSIS AND RESULTS

A total of 7 adults participated, 4 males and 3 females. None had prior experience on radiation emergencies or virtual-reality-based learning. Of them, 42.85% had used virtual reality technology before. 28.5% declared to use virtual reality for leisure. 28.5% had not experienced dizziness in previous experiences. 14.28% knew what radiation emergencies are.

Qualitative Analysis

As previously introduced, a qualitative study of the data was carried out. The study was based on the observations of the participants' behaviour in the recorded videos.

In general, a tendency to perform the same actions in the guided and non-guided parts can be observed. Moreover, all participants try to perform these actions in order, as they were presented to them the first time. The participants showed an improvement between the guided and non-guided parts. In the non-guided part, nothing had to be explained to them. They had learned everything in the guided part.

Analysing the performance of each of the participants separately, it can be observed that previous experience with virtual reality is quite relevant. For instance, participant A used VR every day, and participant B had only used a VR headset once. It could be seen that as familiarity with VR decreases, the duration of the test increases, and the number of interactions and the number of repeated interactions decrease too. Therefore, it can be observed that the average performance is also affected by whether the person uses virtual reality frequently.

Participant A understood the indications of the use of the technology. This participant achieved a similar speed in the guided and non-guided parts, so VR-based interaction was not an impediment for her. This is in line with the questionnaire data. This participant was the only one who managed to complete both stages in the same order. In addition, the number of interactions is the lowest of all participants, mostly avoiding repetition. In the video analysis, the participant would guess correctly how to interact with each of the elements in the environment. For example, the participant anticipated that large objects, unlike small objects, cannot be picked up.

Another illustrative case is participant C. This participant understood the instructions correctly. The results of the non-guided part were like the results participant A. However, participant C did not perform the actions in the correct order. Looking at the order, this participant followed a particular order that, while not properly expected, was also considered valid. As mentioned above, it is important to cut off traffic before going to rescue the injured person, but participant C decided to take this action at the beginning, regardless of his subsequent actions. This participant thought that the traffic had to be cut off first in order to understand the current state of the situation more calmly. On the other hand, participant C spoke to the witnesses at the end of the experience. The participant preferred to find out what had happened to the accident victim and carry out all the actions rather than talk to the witnesses. This shows that for this participant, solving the problem took priority over collecting external information. In relation to the interactions with objects, it was noted that the participant understood how to use them but tried to use them in alternative places to see if something else could be done with them. This leads to an increase in the number of final interactions.

Participant B is the one who had the least relationship with technology, and this had a direct impact on the overall duration of his test. Participant B tried to perform the actions sequentially, protecting the driver before talking to the witnesses. The number of interactions is lower than those for participant C. The total duration was longer. This was perceivable during the experiment as a greater caution when performing actions and interacting with the elements.

In general, the participants adapted to the use of virtual reality technology and managed to complete all the proposed objectives. In addition, they tried to follow the completion order practised in the guided part.

Quantitative Analysis

Along with the qualitative analysis of the participants' performance, a quantitative measurement of the interactions was carried out. The total duration of the tests can be examined in Table 1. In the guided part, the participants had an average duration of 475.43 seconds and in the non-guided part an average duration of 314.42 seconds. A decrease between the guided and non-guided parts can be observed. This indicates an improvement in the average participant performance. It can be considered that the participants have learned to perform the assigned tasks. This skill acquisition may also be influenced by previous experience with virtual reality.

It can be observed that interaction repetition is common. These repetitions usually happened in interactions with elements. In many cases, it is not the participant's intention. For example, when the participant wanted to interact with a small item, and to do so she had to select the item with the controller. Being a small element and given the sensitivity of the controller, even when the participant tried to aim only once, the virtual system would detect that the interaction happened several times, which generates a lot of noise. For this reason, all those interactions that have been performed in the same second on the same element were considered noise repetitions. In order to perform the analysis, and by taking into account the analysis of the videos, this kind of repetitions have been discarded from the samples.

Table 1. Duration of participant tests. This duration is measured in seconds.

Time elapsed for completing the virtual scenario		
	mean (seconds)	**stdev**
Guided	475.43	88.93
Non-guided	314.42	112.38

Table 2. Fulfilled objectives in guided and non-guided stages.

	Fulfilled objectives			
	Number of fulfilled objectives (mean)	**Number of fulfilled objectives (stdev)**	**Time between objectives (mean)**	**Time between objectives (stdev)**
Guided	9	0.0	59.42	11.11
Non-guided	13.57	3.95	27.23	13.32

An increase in the number of interactions in the non-guided part was expected because no repetition of interactions was allowed in the guided part. Therefore, the guided part had the minimum number of possible interactions. However, the time between them is expected to decrease in the non-guided part because the participant is expected to have learned the objectives. The corresponding data is shown in Table 2. This has a direct impact on the final duration of the guided and non-guided parts. It should be noted that all the participants managed to complete all the objectives, which partially evidences the effectiveness of the simulation.

It can be observed that the number of completed objectives is necessarily the same in the guided stage. As mentioned above, repetitions are not allowed (i.e., interactions cannot be carried out more than once). This results in a standard deviation of 0.0. In the non-guided stage, however, the interactions could be repeated.

The expected results with interactions were the opposite. The number of these interactions is expected to decrease as the participant learns which objects to use, how, and when. The participant had also learned what each object is for, so she only interacts with them when it is necessary. Because of this, the time between interactions should increase, as the participant will try to perform the relevant actions only when needed. This approach is observed to be correct given the results obtained (see Table 3).

These data show that the participant acquired knowledge between the guided and the non-guided stages. As it can be seen, duration and number of interactions decreased for all participants. At the same time, they managed to complete all the objectives. Moreover, all of them tried to complete the objectives in the same order as they set out at the beginning, although only one of them succeeded in doing so. However, the remaining participants followed the completion order learnt in the guided stage in most of the non-guided interactions.

Table 3. Means and standard deviations of interactions.

	Interactions			
	Number of interactions (mean)	**Number of interactions (stdev)**	**Time between objectives (mean)**	**Time between objectives (stdev)**
Guided	531.57	134.47	0.88	0.16
Non-guided	291	75.40	0.92	0.19

CONCLUSION AND FUTURE WORK

Training participants in hazardous scenarios involves a series of risks which do not take place in a virtual reality scenario. This paper has reported on a VR-based methodology that assesses the performance increase based on the fulfilled objectives and the interactions with the environment, which is assumed to approximate the expected response in certain scenarios. A pilot study in a virtual reality radiation emergency environment has been conducted. The results indicate coherence between qualitative and quantitative data, and relative improvement in the participant's performance.

As previously introduced, the general applicability of the results is limited by the number of participants. Given the need to increase the volume of samples, future work contemplates carrying out the experiment with a larger number of participants. In addition, the created VR system is intended to study the learning process in radiation emergency situations with groups of professionals dealing with this type of situation. Therefore, it is important to set up an experiment with these profiles (police, doctors, firefighters), which are in principle not expected to have any VR training. According to the presented findings, this could have an impact on the initial performance of the experiences. However, as the training with the VR tool progresses, this deficiency is expected to decrease.

This work is focused on the learning process in potentially stressful situations. Another significant contribution to the work is the study of the participants' stress. For this purpose, the stress of participants in this virtual environment application must be analysed, and it is expected to provide additional data about the different behaviours and interactions of participants during learning. A common alternative for detecting stress and anxiety levels is by means of biometric sensors, which make the experiments less comfortable and realistic for participants. A good balance between detection capabilities and interaction possibilities will be explored.

ACKNOWLEDGMENT

This work has been partially supported by the projects ADARVE (SUBV-20/2021), funded by the Spanish Council of Nuclear Security; CANTOR (PID2019-108927RB-I00), funded by the Spanish Ministry of Science and Innovation; and EA-DIGIFOLK (101086338), funded by the European Commission.

REFERENCES

Aïm, F. *et al.* (2016) 'Effectiveness of virtual reality training in orthopaedic surgery', *Arthroscopy: the journal of arthroscopic & related surgery*, 32(1), pp. 224–232.

Benoit, A. *et al.* (2009) 'Multimodal focus attention and stress detection and feedback in an augmented driver simulator', *Personal and Ubiquitous Computing*, 13(1), pp. 33–41.

Botella, C. *et al.* (2017) 'Recent progress in virtual reality exposure therapy for phobias: a systematic review', *Current psychiatry reports*, 19(7), pp. 1–13.

Bouchard, S. *et al.* (2012) 'Using biofeedback while immersed in a stressful videogame increases the effectiveness of stress management skills in soldiers', *PloS one*, 7(4), p. e36169.
Checa, D., Miguel-Alonso, I. and Bustillo, A. (2021) 'Immersive virtual-reality computer-assembly serious game to enhance autonomous learning', *Virtual reality*, pp. 1–18.
Emmelkamp, P. M., Meyerbröker, K. and Morina, N. (2020) 'Virtual reality therapy in social anxiety disorder', *Current psychiatry reports*, 22(7), pp. 1–9.
Freeman, D. *et al.* (2018) 'Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial', *The Lancet Psychiatry*, 5(8), pp. 625–632.
Freina, L. and Ott, M. (2015) 'A literature review on immersive virtual reality in education: state of the art and perspectives', in *The international scientific conference elearning and software for education*, pp. 10–1007.
Ishaque, S. *et al.* (2020) 'Physiological signal analysis and classification of stress from virtual reality video game', in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, pp. 867–870.
Joëls, M. *et al.* (2006) 'Learning under stress: how does it work?', *Trends in cognitive sciences*, 10(4), pp. 152–158.
Kaplan, A. D. *et al.* (2021) 'The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis', *Human factors*, 63(4), pp. 706–726.
Kavanagh, S. *et al.* (2017) 'A systematic review of virtual reality in education', *Themes in Science and Technology Education*, 10(2), pp. 85–119.
Keitel, A. *et al.* (2011) 'Endocrine and psychological stress responses in a simulated emergency situation', *Psychoneuroendocrinology*, 36(1), pp. 98–108.
Klingner, S. *et al.* (2020) 'Firefighter Virtual Reality Simulation for Personalized Stress Detection', in *German Conference on Artificial Intelligence (Künstliche Intelligenz)*. Springer, pp. 343–347.
Lopez Vazquez, E. (2001) 'Risk perception interactions in stress and coping facing extreme risks', *Environmental Management and Health*, 12(2), pp. 122–133.
Mantovani, F. *et al.* (2003) 'Virtual reality training for health-care professionals', *CyberPsychology & Behavior*, 6(4), pp. 389–395.
Mevlevioğlu, D., Tabirca, S. and Murphy, D. (2022) 'Real-time anxiety prediction in virtual reality therapy: research proposal', in *Proceedings of the 13th ACM Multimedia Systems Conference*, pp. 352–356.
North, M. M. and North, S. M. (2016) 'Virtual reality therapy', in *Computer-assisted and web-based innovations in psychology, special education, and health*. Elsevier, pp. 141–156.
Pallavicini, F., Pepe, A. and Mantovani, F. (2022) 'The Effects of Playing Video Games on Stress, Anxiety, Depression, Loneliness, and Gaming Disorder During the Early Stages of the COVID-19 Pandemic: PRISMA Systematic Review', *Cyberpsychology, Behavior, and Social Networking* [Preprint].
Pine, R. *et al.* (2020) 'The effects of casual videogames on anxiety, depression, stress, and low mood: a systematic review', *Games for health journal*, 9(4), pp. 255–264.
Pluntke, U. *et al.* (2019) 'Evaluation and classification of physical and psychological stress in firefighters using heart rate variability', in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, pp. 2207–2212.

Porcelli, A. J. and Delgado, M. R. (2017) 'Stress and decision making: effects on valuation, learning, and risk-taking', *Current opinion in behavioral sciences*, 14, pp. 33–39.

Porter, A. M. and Goolkasian, P. (2019) 'Video games and stress: How stress appraisals and game content affect cardiovascular and emotion outcomes', *Frontiers in psychology*, 10, p. 967.

Reznek, M., Harter, P. and Krummel, T. (2002) 'Virtual reality and simulation: training the future emergency physician', *Academic Emergency Medicine*, 9(1), pp. 78–87.

Rothbaum, B. O. *et al.* (2000) 'A controlled study of virtual reality exposure therapy for the fear of flying.', *Journal of consulting and Clinical Psychology*, 68(6), p. 1020.

Rudland, J. R., Golding, C. and Wilkinson, T. J. (2020) 'The stress paradox: how stress can be good for learning', *Medical education*, 54(1), pp. 40–45.

Ruesseler, M. *et al.* (2010) 'Simulation training improves ability to manage medical emergencies', *Emergency Medicine Journal*, 27(10), pp. 734–738.

Schwabe, L. *et al.* (2012) 'Stress effects on memory: an update and integration', *Neuroscience & Biobehavioral Reviews*, 36(7), pp. 1740–1749.

Smith, M. J., Conway, F. T. and Karsh, B.-T. (1999) 'Occupational stress in human computer interaction', *Industrial health*, 37(2), pp. 157–173.

Soyka, F. *et al.* (2016) 'Enhancing stress management techniques using virtual reality', in *Proceedings of the ACM symposium on applied perception*, pp. 85–88.

Starcke, K. *et al.* (2008) 'Anticipatory stress influences decision making under explicit risk conditions.', *Behavioral neuroscience*, 122(6), p. 1352.

Subahni, A. R., Xia, L. and Malik, A. S. (2012) 'Association of mental stress with video games', in *2012 4th International Conference on Intelligent and Advanced Systems (ICIAS2012)*. IEEE, pp. 82–85.

Van Wyk, E. and De Villiers, R. (2009) 'Virtual reality training applications for the mining industry', in *Proceedings of the 6th international conference on computer graphics, virtual reality, visualisation and interaction in Africa*, pp. 53–63.

Wemm, S. E. and Wulfert, E. (2017) 'Effects of acute stress on decision making', *Applied psychophysiology and biofeedback*, 42, pp. 1–12.