Virtual Reality (2026) 30:2
https://doi.org/10.1007/s10055-025-01240-y

ORIGINAL ARTICLE

Effectiveness of information modality in virtual reality tutorials

Alejandro Villar[1] · **Carlos León**[1]

Received: 25 April 2025 / Accepted: 15 September 2025

Abstract

Virtual reality can be the optimal trade-off between cost, safety, user experience and learning effectiveness in specific scenarios, like those involving security risks. However, subjects not trained on using virtual systems can have difficulties when using with the virtual environment, preventing them from properly using or understanding the interactions. Immersive tutorials can lower the entry barrier and help users get acquainted with the system, but the learning objectives and interaction methods can be diverse and require different information modalities for a proper understanding. We have conducted a study in which tutorials in a virtual reality system were conveyed with different combinations of text and video with the objective of quantifying the effectiveness of different information modalities. Participants were divided into 4 groups with different information modality combinations, in scenarios for learning how to proceed in radiation emergencies. Results indicate that text effectively conveys new and essential system concepts to users, particularly in systems designed to reinforce a specific concept; and video is more effective when visually demonstrating basic variations of previously learned actions. Additionally, it has been observed that information conveyed through videos promotes interactions with both hands interchangeably. The results can be applied to the design of better adapted tutorials in virtual reality.

Keywords Virtual reality · Tutorials · Learning processes · Information modality · Training · Human-computer interaction

1 Introduction

Virtual reality has been used for training in a wide range of scenarios. Although its relative benefits have been thoroughly evidenced and discussed (Jensen and Konradsen 2018), its use is not always possible or straightforward because users are not always familiarized with the underlying technologies and the interaction processes (Miguel-Alonso et al. 2023). Among these, those involving security risks are especially relevant since virtual reality environments are safer than real experiences and typically imply a lower cost than real simulations (Farra et al. 2019; Naranjo et al. 2020).

An example of a hazardous scenario in which virtual reality training is beneficial is radioactive emergency training (Villar and León 2023b). Radioactive emergencies are those events in which radioactive material is involved, differing from nuclear emergencies in that radioactive emergencies happen outside a nuclear power plant (Isaksson and Raaf 2017). For example, transporting radioactive material or earthquakes that produce a radioactive material leakage into a river, among others. Given the presence of radiation

C. León contributed equally to this work.

This publication is part of the R&D &I projects DARK NITE, PID2023-146308OB-I00, funded by MICIU/AEI/10.13039/501100011033/ and ERDF, EU "A way of making Europe"; ADARVE (SUBV20/2021), funded by the Spanish Council of Nuclear Security; CANTOR (PID2019-108927RB-I00), funded by the Spanish Ministry of Science and Innovation; and EA-DIGIFOLK: An European and Ibero-American approach for the digital collection, analysis and dissemination of folk music project (grant agreement no. 101086338), funded by the European Commission.

✉ Alejandro Villar
avillarrubio@ucm.es

Carlos León
cleon@ucm.es

[1] Department of Software Engineering and Artificial Intelligence, Computer Science Faculty, Universidad Complutense de Madrid, Madrid 28040, Spain

Published online: 27 November 2025

and the variety of emergencies, professionals in the sector must be prepared to act quickly and minimizing mistakes that may adversely affect their and the population's health.

Typically, real life simulations are used for professional training in radioactive emergencies. This entails very high economic and organizational costs. In addition, the use of radioactive material in real training is potentially dangerous. For this reason, radiation rates of this radioactive material are sometimes simulated with a large number of boxes with radioactive labels at different locations on the stage, which indicate the radiation rate at that particular point (IAEA 2005). This limited representation leads to a lesser effectiveness, as the radiation detectors cannot be used correctly. In addition, some situations cannot be recreated, such as emergencies caused by catastrophes like earthquakes or tsunamis or radioactive clouds from an explosion at a nuclear power plant. These differences with the real world mean that training does not achieve full immersion, which limits the correct user performance during the exercise (Makransky and Petersen 2021). These problems can be found in different sectors that have similar training processes. Therefore, many study programs have begun to use other technologies, such as virtual reality, to ensure the effectiveness and safety of training (Gürer et al. 2023; Lele 2013).

However, it is common to observe that trainees who are not used to virtual reality devices find it difficult to familiarize and adapt to the technology (Miguel-Alonso et al. 2024; Rodrigues et al. 2022). In order to achieve good level of immersion and thus better training performance, users must be able to interact with the system naturally and with as little cognitive effort as possible, but specific training may be required (Lackey et al. 2016).

To partially mitigate this situation, virtual reality vendors usually distribute generic training materials such as manuals, interactive tutorials, or little experiences meant as an on-boarding stage for end users (Meta 2019; Valve Corporation 2014). These materials typically cover the first interactions and fundamental concepts of using the virtual reality controllers and basic ways of use found in many virtual reality applications. However, this information is usually quite generic and subject to change depending on the specifications of the application at hand, not covering platform-specific aspects that must be learned for every system or, in the worst case, revisited for each training scenario.

Another approach is to train users on specific systems. Serious video games used for training typically use system-specific tutorials as a learning tool for their mechanics (Checa and Bustillo 2020). The idea is that tutorials are immersive, step-by-step experiences in which the basic information for the user to interact with the system is delivered simply and progressively while interacting according to the same environment mechanics that the user will have to apply during the virtual experience (Villar and León 2023a). How to design system-specific tutorials and their measurable effectiveness has been explored, and a substantial number of studies on improving users' learning have been carried out (Fowler and Cusack 2014; Gee 2005; Ho 2017).

The way in which information is presented to the user must be considered for designing an effective tutorial (Green et al. 2017). Text-based, audio (Rettinger et al. 2023), blocking spatial movement until some criteria have been satisfied, or context-related tutorials (Frommel et al. 2017) are different ways for representing information in virtual tutorials. In a virtual reality environment, while text is a helpful way to explain the feature, tutorial videos can provide immersive visual guidance, allowing users to see and understand actions in a realistic, three-dimensional context. In order to understand the impact that videos can have, it is necessary to compare different information modalities. For instance, a combination of video with explanatory text, with each helping to mitigate the shortcomings of the other, is a plausible way of describing interaction with a possible positive impact on learning. Several studies have explored the use of videos showing the user which actions have to be performed, but their relative improvement over other alternatives has not been explicitly compared (Damen et al. 2014; Goto et al. 2010; Petersen and Stricker 2012; Thoravi Kumaravel et al. 2019). This lack of quantification must be addressed in order to properly inform tutorials in virtual reality environments and provide information in the most adapted and effective modality, depending on the kind of information that is being transmitted.

In this work, the qualitative impact of the use of texts and videos in an interactive tutoring process has been studied. In order to run the study, a virtual reality tool and scenarios focused on radiation emergencies have been developed. Additionally, a learning process through a tutorial has been developed. This tutorial is intended to allow users to understand and observe the actions through texts and videos. The tool is composed of two scenarios with a previous learning phase each. During the tutoring phase, the user should learn how to interact with the elements of the virtual environment and how to perform some basic actions for acting in this kind of emergency. With all these materials, an experiment with 48 users was conducted. Data analysis was carried out using the users' actions gathered within the virtual reality, along with user experience, usability, and cognitive load collected in subsequent questionnaires.

The rest of this paper is structured as follows: section 2 summarizes the previous work on tutorials in virtual reality environments and related studies. Section 3 introduces the impact of information modality on learning effectiveness in virtual tutorials. Section 5 describes the learning platform

and scenarios that have been developed for the experiment, and section 6 describes the experimental setup. Section 7 summarizes and analyzes the gathered results. The discussion and main conclusions and provided in sections 8 and 9 respectively.

2 Previous work

The use of virtual reality to enhance learning processes has increased significantly in recent years (Radhakrishnan et al. 2021). Studies on the application of virtual reality technology in education highlight improvements in cognitive, psychomotor, and affective skills. However, it can sometimes be counterproductive because of specific technological challenges or because the immersive experience distracts from the learning task (Jensen and Konradsen 2018). Virtual environments in higher education are mainly focused on performance and usability, but their widespread application is still in an experimental state (Radianti et al. 2020). However, it has been suggested that virtual reality can potentially influence the primary and secondary students' attendance, knowledge transfer, and skill acquisition, enhancing their attitude toward their learning (Pellas et al. 2019). In the healthcare sector, virtual reality shows similar learning outcomes to standard methods (Sakowitz et al. 2020), but in this case, the patient's experience is improved by giving the medical staff cutting-edge tools for training and skill improvement (Singh et al. 2024). Radhakrishnan et al. focus on studying the use of this technology in the industrial sector and highlight its capability for effective training (Radhakrishnan et al. 2021). Military corps have been using it to improve training skills for many years because it offers a training method without causing any loss or damage to humans and equipment for a manageable price (Lele 2013). The mining sector has studied its use with satisfactory results (Gürer et al. 2023), and promising findings suggest that virtual reality can achieve more effective mental and physical training for emergency situations in the future (Engelbrecht et al. 2019). As Narciso et al. suggest, virtual training provides participants with spatial presence effectively while minimizing cybersickness (Narciso et al. 2020). In addition, this realism also causes physiological alterations in participants, without differences between the virtual reality training and the real-world exercise in terms of the level of stress and task performance (Clifford et al. 2019).

Therefore, virtual environments can help trainees achieve relatively maximal learning outcomes (Nassar et al. 2021), especially engaging and motivating novel users (Sun and Chen 2023). However, the technology can be a barrier for these novel users, even when the environments are specifically designed for them (Ho 2017). The complexity of these platforms and applications negatively affects adaptability and increases the need for a preliminary tutoring phase. A lack of familiarity with the technology can sometimes lead to a difficult adaptation to the new virtual environment. Bad user experience and low usability of the tool can result in ineffective training processes (Checa et al. 2021), sometimes performing worse than traditional methods (Gavish et al. 2015). Because of these usability issues, companies such as Meta (Meta 2019) and Steam (Valve Corporation 2014) are working to continually improve the user experience by adding basic introductory tutorials. The use of tutorials is well-established in the entertainment and video game industry due to the complexity of these systems (Andersen et al. 2012), and, for this reason, tutorials can be a feasible technique for introducing new concepts providing an effective learning process (Villar and León 2023b). Nevertheless, these tutorials often cover generic functionalities that may change based on the specific application's requirements. These applications usually develop their tutorials guided by the limited depth of the ones provided by virtual reality device vendors. Kao et al. found that 90% of video games include a tutorial phase, either at the beginning of the experience or throughout it (Kao et al. 2021). Additionally, research shows that only 10% of virtual reality studies incorporate tutorials to teach users how to interact with the system (Checa and Bustillo 2020). Due to study-specific needs, it is common for developed applications to include a tailored tutorial that teaches users the main interactions and concepts of the test environment (Bhargava et al. 2018; Kleven et al. 2014; Makransky et al. 2019). However, other studies aim to create a general preliminary tutoring phase that covers the core functionalities of virtual reality (Miguel-Alonso et al. 2023).

2.1 Interactive tutorials and information modalities in previous virtual environments tools

The existing literature suggests that the modality in which information is presented significantly influences the effectiveness of learning (Frommel et al. 2017; Kao et al. 2021). Apart from textual format, content can also be presented visually. It is, therefore, essential to consider how information is presented to the user to ensure the effectiveness of the learning process. Green et al. propose three types of tutorials based on: (1) step-by-step instructions that guide the user, (2) examples that demonstrate consequences, and (3) carefully designed experiences that allow users to practice freely (Green et al. 2017). Frommel et al. emphasize that this information should be *context-sensitive*, meaning the tutorial should be an integral part of the experience (Frommel et al. 2017). The provided information must be relevant

to the current task to maintain immersion. Kao et al. note that text is the most common way to present information (88%), and about only half of the examined systems use images (56%). While text produces worse results, it remains the most common across applications, probably due to a lower development cost (Kao et al. 2021). Information can also be presented through visual demonstrations, which have a positive impact and are well-received by users (Thoravi Kumaravel et al. 2019).

During the development of an interactive tutorial, knowing what the users must learn and how the system can be designed to help them is essential. The most basic concepts must be considered to achieve a good learning process in a virtual reality experience. LaValle proposes some fundamental interactions that can be found in a virtual reality experience (LaValle 2023). Among these, *locomotion* refers to how the user can move through the environment. Many ways of movement have been studied, but *point & teleport* has the best results (Bozgeyikli et al. 2016). Teleportation is a relevant concept that users should learn if they cannot move in real life or the available space is limited (Shewaga et al. 2020). The interactions that the user can perform with objects are defined as: *selection*, how the user can detect which objects are interactable; *manipulation*, the action of grabbing and dropping; and *placement*, putting an object in a specific place. *Social interaction* with virtual characters is also relevant and can lead to a better immersion. *User interface interaction* allows users to interact with floating panels and buttons, which can be highly important if the user is stuck and needs to demand help, which is a functionality that improves the general effectiveness (Andersen et al. 2012).

Mayer's Cognitive Multimedia Learning Theory can be used as a foundation to create interactive tutoring processes (Mayer 2014). It provides a list of 12 principles that can be used to enhance learning in desktop applications. Some of the most important principles include the *coherence principle*, which states that all irrelevant material should be avoided; the *personalization principle*, which suggests that information should be presented using an informal narrative through dialogues with virtual characters (Vosinakis 2020); and the *signaling principle*, which emphasizes the importance of highlighting key elements. In this regard, head-up display notifications are the most effective, although they can be intrusive. Another way to prevent users from getting lost is through visual cues (Dillman et al. 2018). These cues can help users identify which objects can be interacted with or where they need to go.

Mayer proposed that knowledge could be classified into *descriptive knowledge* and *explanative knowledge*. Descriptive knowledge refers to knowledge in which the key components or events are described but not explained, such as a list of facts, a description of a series of events, or a set of procedural instructions. In contrast, explanative knowledge refers to knowledge in which the functional relationships among two or more variables are expressed, and the underlying functional rules or mechanisms are explained.

3 Virtual reality platform with different information modalities for training in radiation emergencies

As summarized in section 2, virtual reality can positively influence learning and training processes. However, lack of familiarity with technology and virtual environments is a barrier to effective learning. This study focuses on the importance of a tutoring process by analyzing the impact of different methods to present an explanation in the context of virtual reality training, specifically through the integration of text and video during an interactive tutoring process.

While interactive tutorials have proven to be a helpful tool to remove this barrier, it is important to understand the impact of different aspects of the specific explanations during the process. Among other modalities, these explanations can be presented textually and visually. In this paper, we seek to quantify the effects of information modalities on (1) user performance, (2) user experience, (3) usability, and (4) cognitive load. Knowing the user's behavior and performance is important to understand how effective a training process has been. This performance can be measured through the number of completed objectives, the number and kind of interactions with objects, and movement efficiency, i.e. how much of the user's movement takes her or him to a place where interactions take place.

User experience, system usability, and cognitive load when performing tasks are also important aspects that may affect learning. Our general hypothesis is that while tutorials have in general a positive influence on learning (Villar and León 2023a), the use of text and video has different effects. Video is hypothesized to be an effective way to learn how to carry out actions derived from existing knowledge, due to the visual representation of that action (Yousef et al. 2014). It is assumed to provide a better user experience and performance (Thoravi Kumaravel et al. 2019). Text is hypothesized to be better suited to new, complex actions.

The knowledge categories proposed by Mayer suggest that information modality (text or video) impacts learning differently (Mayer 1995). *Descriptive concepts* can benefit from textual information by being focused on the description of events, instructions, or important data without the need to go into functional details. On the other hand, *explanative concepts* focus on functional relationships. Video can help to visualize the functionalities of the elements interacted

with in the virtual environment. Against this background, we propose the working hypothesis that:

Hypothesis *For both experienced and non-experienced users, information modality has a quantifiable impact in learning in virtual reality environments, where text is better suited for descriptive concepts and video is better suited for explanative concepts.*

More specifically, the hypothesis states that each piece of tutorial information can be more effective if the modality is appropriate, depending on the kind of information being conveyed, and that these differences are observable and statistically significant among subjects. In particular, textual information is better when Mayer's descriptive concepts are being taught, and video is better for Mayer's explanative concepts. The hypothesis motivates the development of the platform described in section 5 and the experiment described in section 6.

4 Usability, user experience and cognitive load

Testing the hypothesis requires assessing not only the learning effectiveness, but also other aspects like usability, user experience and cognitive load. *Usability* has various definitions (Bevana et al. 1991). One of the most widely accepted definitions of usability is the effectiveness, efficiency, and satisfaction with which specific users can achieve specific goals in a particular environment (Brooke 1995). It is important to distinguish usability from *user experience*. Several studies define user experience as a set of characteristics focused on the quality of the product, even taking into account the user's emotions (Hassenzahl and Tractinsky 2006). There is a tendency to say that usability is part of the user experience.

Sometimes, a bad design and approach can lead to incorrect system usability, and the user may have an unsatisfactory experience when using a new tool. Considering the application design in unfamiliar environments, such as virtual reality, is important. Users who are not used to this technology may be affected by sensations of dizziness that hinder the proper execution of tasks (Somrak et al. 2019). Moreover, this study seeks to explore how different ways of presenting information impact the learning process, which can also be affected by a bad user experience (Gordillo et al. 2014).

Due to the importance of these factors, the User Experience Questionnaire (UEQ) has been chosen to measure the user experience (Laugwitz et al. 2008). This method consists of 6 traits that allow the measurement of aspects related to system usability (efficiency, perspicuity, and dependability) and user experience (originality, stimulation)[1]. These traits are defined as follows:

- *Attractiveness.* Overall impression of the product. Do users like or dislike it?
- *Perspicuity.* Is it easy to get familiar with the product and to learn how to use it?
- *Efficiency.* Can users solve their tasks without unnecessary effort? Does it react fast?
- *Dependability.* Does the user feel in control of the interaction? Is it secure and predictable?
- *Stimulation.* Is it exciting and motivating to use the product? Is it fun to use?
- *Novelty.* Is the design of the product creative? Does it catch the interest of users?

It is also relevant to address the *cognitive load* during the learning process (Kirschner 2002). Cognitive load refers to the amount of information that working memory can process at a given moment (Sweller 2011). If the learning process imposes a high cognitive load on the user, it may negatively impact the final results. Various studies indicate that cognitive load can be measured using bio-signals such as EEG, ECG, or EDA, among others (Antonenko et al. 2010). However, their use is highly intrusive, and the measurement of these bio-signals is highly sensitive to any external changes. For example, EEG can be influenced by nearly imperceptible electromagnetic signals emitted by electronic devices, such as those in virtual reality. Studies show that it is possible to use EEG and virtual reality simultaneously by modifying the signals emitted by the headsets (Tauscher et al. 2019). However, the user's movement must be limited. Therefore, questionnaires have been chosen to measure this. In this study, the NASA-TLX questionnaire has been used, consisting of 6 traits that allow the user to assess the system's demands (Hart 2006). These traits and the results obtained are as follows:

- *Mental Demand.* How simple or complex the user perceives the task to be.
- *Physical Demand.* How much physical activity was required. This factor is not decisive because this experience is not designed for high physical activity. Virtual reality for training aims to assist in learning basic concepts that do not require physical effort while using a limited physical space.
- *Temporal Demand.* Analyzes the speed at which elements were presented to the user, that is, the time needed to complete objectives or the pace at which the user

[1] Official *User Experience Questionnaire* web page: https://www.ueq-online.org, accessed on 13/03/2025.

was exposed to stimuli. This factor is not very decisive, as the elements that make up the experience were presented gradually, giving the user the time necessary to complete each objective.

- *Performance.* Measures the user's perception of their success in completing the proposed tasks. Users who have completed the experience are expected to consider themselves more successful. In contrast, users who took a long time to complete certain tasks or who did not complete the experience will have a worse perception.
- *Effort.* Shows how much the user had to exert themselves to complete the task. This factor takes into account both mental and physical effort. It is expected that with less information provided, the more effort it will take for the user. This is because the user has to make extra effort to figure out what they need to do.
- *Frustration.* The user's feeling may be related to the lack of success in completing the tasks. It is expected that users who took longer to complete the tasks would feel more frustrated.

5 Training for interaction in virtual reality and basic usage of tools for radiation emergencies

A virtual reality platform was developed to test the hypotheses detailed in section 3. The system, *Radiation Room*, is a virtual reality tool developed for training first responders in radioactive emergencies. The design of the platform, radiation models, tools, and human profiles and scenarios has been carried out under the supervision of the Spanish Nuclear Safety Council. The functional requirements of the tool and scenarios were set by the Spanish Nuclear Safety Council. The development of the tool underwent several versions and revisions, during which radioactive emergency professionals participated in some pilot tests, providing their experience and assessment and valuable feedback to the system. Observations of their behavior while using the virtual reality system, complemented by follow-up meetings in which they provided thorough feedback on the current state of the tool and suggestions for essential features to include in the final version, significantly contributed to the evolution of the application. Their participation was therefore crucial for content creation, the effective development of the virtual environment, and the validation of the usefulness of an interactive tutorial phase (Villar and León 2023a). The platform has been developed with Unreal Engine (Epic Games 2025). HTC Vive Pro 2 and Meta Quest 3 headsets were used run the experiments. A frame rate of 120 Hz was used, as it is the maximum supported by the Meta Quest 3 and aligns with the threshold at which users tend to experience reduced symptoms of simulation sickness without a significant negative impact on their overall experience (Wang et al. 2023). The computers used exceeded the specifications recommended by the virtual reality vendors, ensuring optimal performance and stability. As a result, the system is compatible with a range of virtual reality devices and can be adapted to different hardware configurations.

Radiation Room includes two tutorial scenarios and two test scenarios. In the tutorial scenarios, a number of explanations and guides are given. The first scenario (*VR Tutorial*, detailed in section 5.1), focuses on the basic actions to perform within a virtual reality application, and the second tutorial (*Emergency Tutorial*, detailed in section 5.3) is focused on the actions the user must learn to handle a simple radiation emergency. The list of explanations is provided in Table 1. These explanations follow the same pattern in both tutorials.

The two test scenarios of *Radiation Room* (see Fig. 1) are the virtual reality mechanics, *VR Training*; and a basic radioactive emergency, *Emergency Training*. In both test scenarios, some objectives based on the previous tutorial phases must be fulfilled (see sections 5.2 and 5.4). These test scenarios are meant to measure users' performance after the tutorials and measure the performance difference between groups that are given tutorial information with different modalities.

Information panels located in the test scenarios display a list of tasks to be completed (see Fig. 2). These panels provide a clear and organized view of the uncompleted tasks, which is essential to guide the user in the correct execution of his objectives. The task list divides the overall process into smaller, more manageable tasks, which are presented in a logical and sequential order that facilitates the understanding and work flow necessary to achieve the main objectives set. In addition, the visual presentation of the task list on the dashboard acts as a guide and a constant feedback system that helps the user to evaluate their progress.

The layout of the elements and the user interface in the tutorial has been designed based on Mayer's Cognitive Theory of Multimedia Learning (Mayer 2014). According to this, the next principles have been taken into account for the design:

- *Coherence principle.* Irrelevant information should be avoided, such as background sounds or decorative elements. Any distraction may create noise in their learning process. In this way, there are no elements in the environment that cannot be interacted with.
- *Signaling principle.* Visual cues can be used to attract the user's attention. This way, the user can identify important places or objects related to the task they have to perform. This principle has been widely studied (Albus

Table 1 Identifiers for each explanation provided in tutorial scenarios. *VR Tutorial* is the first scenario related to virtual reality actions. The explanations with the identifier *TUT1* belong to this tutorial. *Emergency Tutorial* is the second scenario, where basic concepts of radiological emergencies are explained. *TUT2* is the identifier assigned for the explanations found in it

Id	Tutorial Id	Description
TUT1-MOV	VR Tutorial	Related to the *locomotion* functionality. It explains how to use the teleport system to move around the scenario
TUT1-SEL	VR Tutorial	Related to the *select* functionality. It explains how to detect interactable objects
TUT1-MAN	VR Tutorial	Related to the *manipulation* functionality. It explains how to grab and release an interactable object
TUT1-PLA	VR Tutorial	Related to the *placement* functionality. It explains how to place an object in a specific place
TUT1-SIN	VR Tutorial	Related to the *social interaction* functionality. It explains how to start and end a conversation with an NPC
TUT2-CBR	Emergency Tutorial	Related to the *CBR vehicle* functionality. It explains what a CBR vehicle is and how to get elements and devices from it
TUT2-WAL	Emergency Tutorial	Related to the *walkie-talkie* functionality. It explains how to report the situation to the central office using a walkie-talkie
TUT2-SAB	Emergency Tutorial	Related to the *SABG-100* functionality. It explains what an SABG-100 device is and how to measure the radiation dose with it
TUT2-EPD	Emergency Tutorial	Related to the *EPD* functionality. It explains what an EPD device is and how to measure the radiation dose with it
TUT2-SOS	Emergency Tutorial	Related to the *help panel* functionality. It explains how to demand specific information using the help panel

et al. 2021; Dillman et al. 2018) and has been carefully considered during the design of these tutorials. Signaling has been applied in: (1) the path the user must follow to move on to the next phase of the tutorial, which is marked with arrows on the floor; (2) if the user attempts to interact with an object, it signals when interaction is possible by making the user's hand vibrate and change color (Ghosh et al. 2018); and (3) it has been used to indicate to the user where they have to place the required objects.

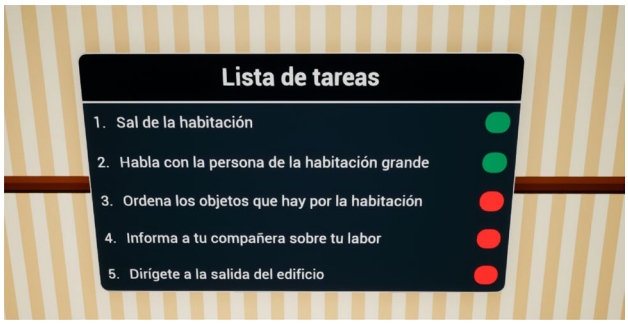

Fig. 2 Panel displaying the tasks list that the user must complete. The panel shown corresponds to the first emergency scenario

- *Redundancy principle.* A complex system can overwhelm the user. Therefore, it is important to minimize any redundant information. For this reason, efforts have been made to ensure that both text and video provide complementary information rather than repeating the same content.
- *Spatial contiguity principle.* Related elements should be placed close together so that the user can understand their relationship. For example, the explanatory text and video should be together, or the explanations and interactive elements should be in the same area (Fig. 3).
- *Temporal contiguity principle.* Presenting textual and visual information about a specific action simultaneously is essential. Splitting them may negatively impact the user by creating confusion about the connection between the two explanations.

Fig. 1 First emergency scenario, *VR Training*, is depicted in the left image. Second emergency scenario, *Emergency Training*, is shown in the right image

- *Segmenting principle.* Learning is more effective when the user receives segmented information rather than all at once. Long and complex functionalities should be broken down into smaller explanations so the user can better understand each part. In this way, the tutorials have been divided into small explanations to focus the user's attention.
- *Multimedia principle.* Learning is improved when text and images are used together instead of just text. It allows users to connect textual explanations with the actions visible in the images. With this in mind, this study aims to analyze the impact of videos instead of images.
- *Personalization principle.* Using formal language can negatively impact learning. It is better to speak in the first person, focusing on the user. Therefore, the virtual tutor uses informal language, aiming to be more of a companion than a teacher.

All elements that could cause distractions have been intentionally avoided in the design of the tutorials (*coherence principle*). As mentioned, each tutorial scenario is explained with a fixed number of five explanations (*segmenting principle*), and every area in which explanations are provided has the same layout (see Fig. 3). Upon reaching the zone, the user can see a central area where the elements she or he needs to interact with are clearly visible. The information is presented on a wall in the back, formatted according to the tutorial modality. A small floating panel also shows the task the user must complete in order to finish the explanation. Finally, a small floating robot can be seen next to this task panel. This robot guides the user through the different stages of the experience, providing context for what the user needs to do in the scenario (Vosinakis 2020).

The details of the two tutorial scenarios are provided in sections 5.1 and 5.3; and the details of the test scenarios are provided in sections 5.2 and 5.4.

5.1 VR Tutorial: basic interactions virtual reality

The lack of familiarity with the technology and the need to learn specific concepts about these emergencies make it necessary to consider a previous step-by-step tutoring process. The first tutorial, *VR Tutorial*, tries to train users about the basic features found in the virtual environment for them to be able to tackle the rest of the process. Interaction with the environment is one of the main aspects of virtual reality, and getting the user to interact with the virtual world as closely as possible to the real world is necessary to obtain a good user experience and a correct performance of the current task. Moving around the environment, grabbing and dropping objects, or interacting with virtual people are some actions to consider when creating a virtual reality tool. LaValle raises these questions and divides these more basic interactions on which this first tutorial scenario is based (LaValle 2023).

For reference, the described actions from this point on are accompanied with their corresponding *id* tag in the experiments. For instance, the tag used for counting the frequency of locomotion actions in the first tutorial, described next, is assigned the *id = TUT1-MOV*.

5.1.1 Locomotion

In many virtual reality scenarios, the virtual world's size is much larger than the delimited zone in the real world. For example, *Radiation Room*'s second scenario takes place in a town (see section 5.4) but the experiments have been conducted in a limited-size room. The user movement inside the virtual world is called *locomotion*.

There are several ways to perform this movement in virtual reality (Boletsis and Chasanidou 2022; Cherni et al. 2020). Some of these techniques are *walk-in-place*, where the user walks in place without moving; *redirected walking*,

Fig. 3 Example of tutorial scenario. The figure represents the *detection of interactable objects* step (*TUT1-SEL*). The description panel can be seen in **A**. The task to be performed by the user is shown in **B**. **C** is the zone where the elements that the user must interact with to complete the task are located. **D** is the wall that blocks the user until the task is completed

where the environment and the scrolling speed are modified so that the user does not leave the bounded area; using the controller's *joystick* for smooth scrolling; or the teleportation technique, where the user must point a laser at the place in the world where they want to be and they will automatically teleport to that position. Bozgeyikli et al. call this way of movement as *point & teleport* and emphasize that it is the most intuitive, easy-to-use, and user-preferred technique (Bozgeyikli et al. 2016). This locomotion technique induces low levels of motion sickness, similar to the presented methods, and is expected to yield particularly positive results in dynamic virtual environments. The *point & teleport* method was chosen for movement in the virtual reality system due to its favorable results. This action is the first tutorial explanation (*id = TUT1-MOV*).

5.1.2 Manipulation

Object *manipulation* is one of the most important user interactions to achieve a good level of realism in a virtual reality system. Humans constantly manipulate objects defined by specific shapes, weight, or functionalities. Implementing the diversity of objects and how we interact with them is a challenge in virtual environments. However, regardless of the type of object the user wants to interact with, the interactions that can be performed with the world can be summarized as follows:

- *Selection (id = TUT1-SEL)*. In the real world, we select the objects we want to interact with unconsciously since we know which objects we can manipulate and how. However, virtual reality environments are complex, and sometimes, there are objects that users cannot interact with. In the used tool, vibration and color change of the virtual hand when approaching the object inform users of the possibility of interacting with those elements.
- *Manipulation (id = TUT1-MAN)*. The most basic interaction that humans can perform with objects is *grabbing* them. This action can also be performed within the virtual world and thus take the objects to particular locations. By grabbing the object, the user can move and rotate it as desired. For example, the user can bring it close to his face and rotate and inspect it. To do this, the user must bring the virtual hand close to the object and then press a controller button, which must be held down to pick it up. Subsequently, the user can also perform the *release* action, releasing the button, to drop the object anywhere.
- *Placement (id = TUT1-PLA)*. Users can drop objects anywhere, but some specific tasks may require them to place objects in specific locations. For example, placing traffic cones at specific locations on the road. When the user picks up an object, a circle appears on the floor with small representations of that object floating around. Then, when the user drops it in a place near the indicated location, the object will be automatically positioned at the target location. This interaction facility is called the *basin of attraction* concept.

This list of actions shows the main interactions that can be performed with most objects found in real life and a virtual environment. However, objects often have specific functionalities, allowing users to perform different actions. For example, using a *walkie-talkie* to report some information or using the *dosimeters* to measure the radiation of an object.

5.1.3 Social interaction

In the virtual reality scenarios, social interactions with other users or with characters created specifically for some purpose can take place. These characters, not controlled by any user, are called Non-Playable Characters (NPCs), and are typically essential for a realistic experience to be achieved. For example, witnesses and victims in emergencies can be NPCs. The user should be able to interact with these characters as they would with a real person to achieve a realistic system. In the developed virtual scenarios, users can engage in conversation with these characters, understand the current situation better, and sometimes, find out the next task to be performed. To start a conversation with an NPC, the user must get close to her or him and interact with a floating button that appears in front of the NPC with the phrase "Click to talk" in it. Also, if the user moves away, the character will automatically stop talking. How to trigger a conversation with an NPC is the last explanation that appears in the first tutorial (*id = TUT1-SIN*).

5.1.4 Interaction with user interface elements and characters

The virtual reality platform also includes user interface elements. These interfaces are based on panels with buttons, informative texts, or boxes to enter text. In a virtual world, these panels are presented as floating objects. In virtual reality, interactions are performed using controllers attached to each user's hand. In most cases, the user's controller casts a visible beam forward, allowing users to interact with whatever the beam touches (see Fig. 4). For example, this action is used to start conversations with NPCs, as explained above.

Information on the specific mechanics of basic user interface elements was not explicitly given in the tutorial. Users need to know this action beforehand to watch the video and activate the task, and if the user does not know how to press the button to start the task, the rest of the experiment is not

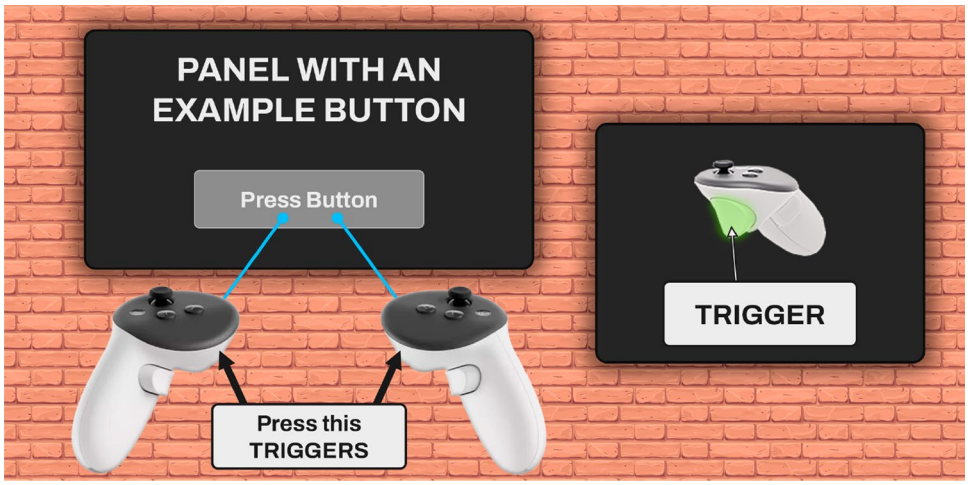

Fig. 4 A depiction of an interaction with an interface shown to the user at the beginning of the experiment. Pressing buttons on the panels is carried out by pointing with the controller and pressing the front trigger

Table 2 List of tasks to be performed by the user in the first practice scenario, *VR Training*. The identifier indicates the scenario in which it is performed, *VR*. The column *Tutorial Learnings* shows which tutorial explanations are required in order to perform this task

Id	Tutorial learnings	Task
VR1	TUT1-MOV	Exit the small room
VR2	TUT1-MOV, TUT1-SIN	Talk with the person in the large room
VR3	TUT1-MOV, TUT1-SEL, TUT1-MAN, TUT1-PLA	Put the objects in their correct places
VR4	TUT1-MOV, TUT1-SIN	Report the person that you have completed the work
VR5	TUT1-MOV	Go to the exit of the building

possible. Adding this explanation in the interactive tutorial means that groups with text explanations would learn how to interact with panels, and they would be able to activate the video or the task to be performed (see section 6.2), which would invalidate the results of those groups.

5.2 VR training: virtual reality mechanics scenario

This first scenario, *VR Training*, follows the first tutorial, *VR Tutorial*, and it has been specifically designed to familiarize the user with the basic interactions within a virtual reality environment. The main purpose of this introductory scenario is to serve as a bridge between the explanations presented in the previous tutorial (see section 5.1) and the practical application in more complex situations that will be addressed in later scenarios.

The scenario is composed of three rooms, each separated by doors that remain locked until specific conditions are met. The experience begins in the first room, where a guide robot introduces the user to these training scenarios. The user is informed about the task panel displaying the objectives that must be completed to progress through the scenario (see Fig. 2). Once the introduction is complete, the door opens automatically, allowing access to the second room where the user can begin completing the tasks listed in Table 2. The second room has an NPC and various interactive objects in it. Upon speaking with the NPC, the user knows that their task is to organize objects scattered around the room. Specifically, the user must correctly place traffic cones, and radiation emergency tools like SABG-100 probes or EPD devices in their designated locations. After completing the task, the user must report back to the NPC and then speak to the robot once more. After this interaction, the final door will open, allowing the user to enter the last room and complete the experience. Following the *context-sensitive principle*, both the NPC's dialogue and the objects within the scenario are thematically related to radioactive emergency response.

The tasks and interactive elements within this scenario are designed to allow the user to move through the environment using the teleportation system, talk to characters in order to know the problem they must face, and interact with objects by placing them in specific locations. This setup enables the user to apply the concepts learned during the previous tutorial within the context of a radioactive emergency. Throughout these two tasks, the user should effectively integrate all the actions and knowledge the user has learned in the previous tutorial.

5.3 Emergency tutorial: basic concepts for radiation emergencies

The second tutorial, *Emergency Tutorial*, seeks to teach users how to report the current situation to the head office and measure the radiation with devices used in real radiation emergencies. In this second tutorial, the user has access to a new functionality developed to provide information on demand if the user does not remember a specific interaction. The user will learn the following elements, concepts, and actions:

1. *CBR vehicle (id = TUT2-CBR)*. A CBR vehicle (Chemical, Biological, and Radiological) is a specialized mobile unit equipped with all the necessary materials to act in a basic emergency involving such risks. The interaction with it is carried out through a panel where the user can select the object and the quantities to be extracted. In addition, there is a visual representation and a brief description of the object.
2. *Use walkie-talkie for reports (id = TUT2-WAL)*. The user must report the situation to the head office during radioactive emergency. The initial report is expected to explain the situation at the time of arrival at the emergency. The final report informs about the actions taken and the final status of the scenario. For this purpose, the users can use virtual walkie-talkies to communicate. When grabbing the walkie-talkie, the user is presented a panel where the central office will ask some questions. The user can choose 1 out of 3 possible answers for each question.
3. *SABG-100 interaction (id = TUT2-SAB)*. The SABG-100 is a probe capable of measuring the surface radiation from objects and people, providing the current dose rate of the element under study (the current radiation dose is the quantity of radiation an object emits in a certain time). The measurement is performed by bringing the lower part of the probe within 5 cms of the object to be observed.
4. *EPD interaction (id = TUT2-EPD)*. An Electronic Personal Dosimeter (EPD) can provide a continuous reading of the accumulated dose and the current dose rate (accumulated radiation dose is the total radiation dose accumulated over a defined period such as minute, hour or year). Professionals must carry it close during a radioactive emergency to know the radiation rate they are exposed to.
5. *Help panel (id = TUT2-SOS)*. The help panel allows the user to ask for explanations on demand (Andersen et al. 2012). This way, the user will always be able to see the same explanations shown during the tutorials.

5.4 Emergency training: basic radioactive emergency scenario

There is a wide variety of possible scenarios classified as radioactive emergencies. For this study, it has been decided to recreate an emergency involving the robbery of radioactive material and a car accident. The accident takes place in a town near a gas station. In this scenario, *Emergency Training*, there is an element that emits a low radiation, an injured person, and a witness. In addition, there is a CBR vehicle with all the necessary items: walkie-talkies, traffic cones, SABG-100 probes, and EPD devices (Fig. 5).

The user should gather information about the incident, secure the area, address the problem considering the radiation environment, and report what happened (see Table 3). The objectives list the basic actions to be performed by the professionals involved, with operational protocols in mind. Most of these objectives have a natural order of accomplishment, and the user should follow it. However, the user is free to perform them in any order deemed appropriate, bearing in mind that the order proposed would be the correct one for this scenario.

This scenario is the only one that can be finished without completing all the proposed objectives. According to radioactive emergency response protocols, reporting the final situation using the walkie-talkie is the last step the professional must complete. Therefore, when the user makes the final report, it is considered that they have decided there are no more tasks to be performed.

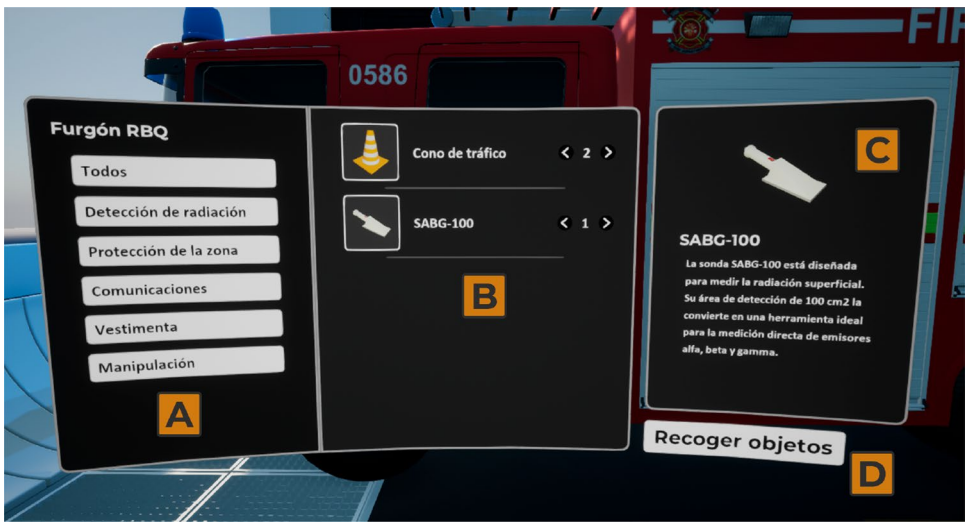

Fig. 5 CBR vehicle interactive panel. The user can filter by object type in (A) panel. (B) shows the object list, allowing the user to select how many objects she/he wants to extract. The user can see a visual representation and description of the object in panel (C). When the user presses the button (D), selected objects will spawn near them

Table 3 List of tasks to be performed by the user in the second practice scenario, *Emergency Training*. The identifier indicates the scenario in which it is performed, *EME*. The column *Tutorial Learnings* shows which tutorial explanations are required in order to perform this task

Id	Tutorial Learnings	Task
EME1	TUT1-MOV, TUT1-SEL, TUT1-MAN, TUT2-CBR, TUT2-WAL	Use a walkie-talkie for the initial report
EME2	TUT1-MOV, TUT1-SEL, TUT1-MAN, TUT2-CBR, TUT2-EPD	Carry an EPD with you
EME3	TUT1-MOV, TUT1-SIN	Talk to the witness
EME4	TUT1-MOV, TUT1-SIN	Move the driver to a safe area
EME5	TUT1-MOV, TUT1-SIN	Obtain information from the injured person
EME6	TUT1-MOV, TUT1-SEL, TUT1-MAN, TUT1-PLA, TUT2-CBR	Use traffic cones to cut traffic
EME7	TUT1-MOV, TUT1-SEL, TUT1-MAN, TUT2-CBR, TUT2-SAB	Measure the radiation of the injured person
EME8	TUT1-MOV, TUT1-SEL, TUT1-MAN, TUT2-CBR, TUT2-SAB	Measure the radiation of the conveyed object
EME9	TUT1-MOV, TUT1-SEL, TUT1-MAN, TUT2-CBR, TUT2-WAL	Use a walkie-talkie for the final report

6 Experimental setup

Section 3 hypothesized that text is better suited for descriptive concepts and video is better suited for explanatory concepts. This section describes the experimental design carried out to test it. In the experiment, four combined modalities are studied and compared: (1) without information, (2) visual modality, (3) textual modality, and (4) both textual and visual modalities displayed at the same time.

For the experiment, a specific set of hardware and software was carefully selected to maximize efficiency and ensure the validity of the obtained results. The virtual reality devices used were Meta Quest 3 headsets connected to the computers via cable. Although this device can be used wirelessly, a wired connection was chosen to ensure stability and performance during the experimental sessions. The computers used had technical specifications exceeding those recommended by Meta. This ensured they could handle the graphical and computational demands of the experiment without issues, providing a smooth and uninterrupted environment. The testing area measured 3x2 m, which was suitable for allowing the necessary freedom of movement within the virtual reality simulations without compromising safety or data accuracy (see Fig. 6).

6.1 Participants and information modality division

A total of 48 users participated in this experiment. The participants were divided into four groups of 12 users based on the information modality displayed. The groups were designed to keep homogeneity in user profiles (see Table 4).

Each group experimented different tutorial panels with information. The first group, *BaseGroup*, received no explanations. The second group, *VideoGroup*, could only obtain information through video. The third group, *TextGroup*, received explanations in text format. Lastly, *VideoTextGroup*

Fig. 6 Three experiment zones with the same hardware and software characteristics were set up. This allowed multiple tests to be conducted at the same time. Users with virtual reality devices, computers, and delimited zones can be observed in this figure

Table 4 Distribution of participants across the four groups. User profile information is provided, including the average age and gender distribution of each group

	Number of participants	Age		Gender	
Groups		M	SD	Female	Male
BaseGroup	12	43.83	13.1	41.67%	58.33%
VideoGroup	12	35.67	9.85	41.67%	58.33%
TextGroup	12	41.33	8.25	16.67%	83.33%
VideoTextGroup	12	40.25	11.36	25%	75%

had access to the full explanation through both text and video (see Fig. 7).

6.2 Procedure

Before starting the experiment, participants were informed about the study's objective, and an *informed consent* document was provided. This document clearly stated that participants could withdraw from the study at any time without consequences and ensured the anonymous and confidential handling of their personal data. Subsequently, users were given a document containing the basic information necessary to complete the experiment correctly. This document explains how to interact with interfaces (see section 5.1.4). They were also informed that reading the text or watching the video was required before starting the assigned task in each tutorial phase. This part of the document depended on the group to which the participant belonged. For example, *TextGroup* had to read the text, while *VideoGroup* had to watch the video. Once all the information had been reviewed, participants could ask any questions they needed before starting the experiment.

The experiment was divided into two phases: (1) an interaction phase within the virtual reality environment and (2) a questionnaire phase. The interaction phase in the virtual environment followed a well-defined sequence. First, all participants had to complete the initial tutorial (*VR Tutorial*), which focused on familiarizing them with the basic mechanics of virtual reality interactions. After finishing this tutorial, they continued with the first practice scenario (*VR Training*), designed to reinforce the skills acquired in *VR Tutorial*. Then, participants moved on to the second tutorial (*Emergency Tutorial*), which covered specific functionalities for responding to radioactive emergencies. Once they completed this tutorial, they were exposed to the second practice scenario (*Emergency Training*), where they had to apply their knowledge to manage situations involving radioactive sources. No time limit was imposed at any stage, allowing each participant to progress at their own pace and ensuring a proper understanding of the content.

After completing the four scenarios, participants proceeded to the questionnaire phase. In this stage, they were asked to provide personal information, including age, gender, educational level, and other relevant aspects. Subsequently, they answered a set of specific questionnaires regarding the application's usability and user experience (Laugwitz et al. 2008). Finally, they filled a questionnaire designed to measure cognitive load (Hart 2006) experienced during the use of the application. This is detailed in section 7.2.

7 Results and analysis

The data collected can be divided into two categories: results on user performance along the different virtual reality tasks and scenarios, like the time it takes for the user to complete the objectives or the number of interactions with objects (section 7.1); and results on usability, user experience, and cognitive load (section 7.2).

Fig. 7 Tutorial display where the virtual interaction is depicted. The example corresponds to the *Detection of interactable objects* explanation. (A) is the format used for the *BaseGroup*, only the panel title is displayed. (B) shows the explanation by text to *TextGroup* and (C) depicts *VideoGroup*, which received all the information only on video. (D) is represents *VideoTextGroup*, where text and video were used to display the information

Table 5 Overall results about the user's performance in the virtual reality experience. This information is based on the total duration of the experience and the objectives completed. Values are presented as mean (standard deviation) for each group. For each factor, t-statistics, degrees of freedom, and p-values are reported

Groups	Total duration	No. of completed tasks	Time per task
BaseGroup	2376.6 (225)	16.27 (6.62)	146.82 (57.36)
VideoGroup	2473.1 (273.24)	22 (3.7)	110.88 (30.03)
TextGroup	2485.6 (723.99)	24 (0)	89.84 (26.56)
VideoTextGroup	2864.2 (612.24)	22.75 (2.49)	107.47 (28.03)
Degrees of freedom	22	22	22
t-statistics	1.324	−1.667	1.515
p-value	0.049*	0.299	0.106

Table 6 Kruskal–Wallis test results comparing the four groups on overall results about the user's performance in the virtual reality experience

Groups	Total duration	No. of completed tasks	Time per task
BaseGroup vs. VideoGroup	0.7221	0.1182	0.4143
BaseGroup vs. TextGroup	1	0.0029*	0.0166*
BaseGroup vs. VideoTextGroup	0.077	0.0431*	0.2516
VideoGroup vs. TextGroup	0.8506	0.1424	0.1601
VideoGroup vs. VideoTextGroup	0.2021	0.9584	0.9957
TextGroup vs. VideoTextGroup	0.2021	0.2695	0.3692
Total	0.0546*	0.001*	0.0128*

In order to correctly analyze the data collected, different statistical tests were performed to evaluate the differences between groups (Dixon and Massey 1951). First, the *Shapiro-Wilk* normality test was applied to determine whether the variables followed a normal distribution. Since the variables did not meet the normality principle ($p < 0.05$), we opted to perform nonparametric tests. Specifically, we performed the *Kruskal-Wallis* and *Wilcoxon tests*. Subsequently, in order to understand the behavior of the variables analyzed, a descriptive analysis of the data was carried out, including *means (M)* and *standard deviations (SD)*. Although the results of the nonparametric tests have not been included in this paper, the variables highlighted in the following analysis are those that showed statistically significant differences ($p < 0.05$).

7.1 Virtual reality performance

The metrics for user performance are based on the analysis of the actions performed by the user. Interaction events such as grabbing or releasing an object, which object they interact with, when they move, the final position of that movement, or when they complete a task are all stored in a database. These data are then aggregated into more general indicators, as shown next.

7.1.1 Tasks completed and total duration

The total duration of the experience, number of completed tasks, and the time the user needs to complete those tasks allow a first overall analysis of the user's results (see Table 5 and 6). The time taken to complete the objectives is supposed to be affected by the modality in which the information is presented. Therefore, the experience duration is analyzed regardless of the time the user takes to view the videos or read the text.

The total duration represents the time, in seconds, taken to complete all the scenarios. It can be observed that *VideoTextGroup* is the group that has taken the longest to complete the experience ($m = 2864.2, sd = 612.24$), while *BaseGroup* is the one that has taken the least ($m = 2376.59, sd = 225$). Regarding the number of completed objectives, some users have not completed all the proposed objectives. *BaseGroup* completed the least number of objectives ($m = 16.27, sd = 6.62$). The total time the user needs to complete the experience is affected, as it can be observed in a shorter total duration of *BaseGroup*. The rest of the groups managed to complete all objectives in most cases. Considering this, it can be seen how *VideoGroup* ($m = 2473.1$) and *TextGroup* ($m = 2485.6$) have a similar duration, indicating that the time spent by the user watching the video is similar to the time spent reading the text. However, *VideoTextGroup* has a significantly higher average duration than these two groups. This is because the user needs more time to do both, read the text and watch the video. In terms of time efficiency, it can be concluded that the best option is to have only one of the two modalities, as they allow all objectives to be met in less time.

Due to the *BaseGroup* users' failure to complete all the proposed objectives, the average time taken to complete each task was analyzed. In this case, the *BaseGroup* is also the worst-performing group ($m = 146.82, sd = 57.35$). The *TextGroup* takes the least average time to accomplish a goal ($m = 89.84, sd = 26.56$). Knowing the number of completed objectives and the average time to complete each of them, the effective time can be analyzed. With it, it is possible to calculate how users in this group were not very efficient due to re-watching explanations, repeating dialogues, contemplating the environment, or interacting unnecessarily with objects. This states that *VideoGroup* is the group with the highest percentage of effective time, with 94.51%, followed by *TextGroup* with 86.74%.

Table 7 Time taken by users to complete the tutoring phases. The total time, the time for each scenario, and the average time spent on each task are shown. Values are presented as Mean (Standard Deviation)

Groups	Total duration	*VR Tutorial* duration	*Emergency Tutorial* duration	Mean time per task
BaseGroup	915.6 (415.62)	636.08 (405.45)	279.52 (226.44)	134.06 (49.8)
VideoGroup	669.8 (185.95)	449.02 (171.28)	220.78 (61.77)	67.57 (18.76)
TextGroup	581.89 (235.61)	338.96 (155.39)	242.93 (96.19)	58.42 (22.25)
VideoTextGroup	595.87 (223.51)	352.67 (163.8)	243.2 (96.2)	59.07 (22.57)

7.1.2 Tutorial tasks

Performed tasks in tutorial scenarios provide information about the effect of the modality during the learning of each task. The observed time is measured from when the user activates the task to its completion. When activating the task, the user claims they have already read the text or watched the video, so the variation in time is exclusively influenced by the time it takes to complete the task. In this phase of the study, we will analyze exclusively the time it took the user to complete the proposed tasks.

A general analysis of the results shows a remarkable difference between *BaseGroup* ($m = 915.6, sd = 415.62$) and the rest (see Table 7). This happens mainly because certain tasks sometimes require a long time to complete them. *VideoGroup* took the second longest time to complete all the proposed tasks in the tutorial ($m = 669.8, sd = 185.85$). However, the difference compared to other groups is not very significant. This preliminary comparison shows the importance of informing the user regardless of the information modality.

According to the average time taken by each group to complete the tutorials, the concepts explained in *VR Tutorial* are the main problem (see Table 7). The first scenario exposes the user to her/his first interaction with the virtual world, expecting this outcome. Novelty and unfamiliarity with the basic mechanics seem to influence the time the user takes to complete the tasks in this first tutorial. In addition, no significant difference in the time taken by each group to complete *Emergency Tutorial* is observed. It can be deduced that the first minutes and the lack of initial knowledge are the main problems for the users. Once the basics of the system are known, the user only has to acquire related concepts based on these basics.

The average time to complete each task shows a notable difference between *BaseGroup* ($m = 134.06, sd = 49.8$) and the others, almost tripling the average time required to complete each task (see Table 7). This is mainly because some tasks are very time-consuming. *VideoGroup* took the second longest average time to complete each task ($m = 67.57, sd = 18.76$). The main difference relies on the beginning, as the first tutorial shows (see Tables 8, 9 ,11, and Fig. 8).

Table 8 Results related to the time taken by the user to complete the tasks proposed in the first tutorial (TUT1). These tasks focus on movement (MOV), detection of interactable objects (SEL), manipulation of those objects (MAN), placing those objects in specific locations (PLA) and starting a social interaction with a character (SIN). Values are presented as mean (standard deviation) for each group. For each factor, t-statistics, degrees of freedom, and p-values are reported

Groups	TUT1-MOV	TUT1-SEL	TUT1-MAN	TUT1-PLA	TUT1-SIN
BaseGroup	303.36 (218.44)	146.84 (139.19)	55.39 (60.11)	114.50 (103.31)	34.58 (8.49)
VideoGroup	255.76 (134.98)	56.65 (35.84)	31.65 (15.42)	74.06 (55.46)	30.90 (7.97)
TextGroup	75.75 (50.56)	67.13 (53.43)	50.74 (40.87)	111.56 (92.42)	33.78 (7.81)
VideoTextGroup	81.11 (53.48)	88.16 (44.49)	63.39 (41.10)	83.78 (57.07)	36.23 (15.75)
Degrees of freedom	22	22	22	22	22
t-statistics	0.242	1.003	0.724	−0.848	0.461
p-value	0.811	0.327	0.477	0.405	0.649

The first task in the tutorial is movement through teleportation (*TUT1-MOV*). A significant difference can be observed depending on the group based on the existence of text. For example, the time needed by *BaseGroup* ($m = 303.36, sd = 218.44$) is four times longer than the time needed by *TextGroup* ($m = 75.75, sd = 50.56$). It is also observed that *TextGroup* performs the best over *VideoTextGroup* ($m = 81.11, sd = 53.48$). This may state that the presence of video impacts negatively on time. However, the video option is better than the absence of information, as it shows by groups without text (see Table 8 and 9). This may be related to the text and video formats, as the text explains how to perform the interaction and videos show it. At the end of the experiment, users in the groups without text commented that the did not know there were more buttons on the controller, so they did not understand that they had to touch something different in order to move. Once they realized there was another button, their perception of the controller was different, and the following tasks were easier.

In the second task (*TUT1-SEL*), the user must learn to select objects. The results show that *VideoGroup* obtains the best results ($m = 56.65, sd = 25.84$). It seems that the use of text and video has a negative impact on the results. *BaseGroup* took the longest time to complete such a task ($m = 146.84, sd = 139.19$). Observations of user behavior show that they had no problems understanding the functionality. However, users are not used to moving using

Table 9 *Kruskal–Wallis* test results comparing the four groups on the time taken to complete the tasks in the first tutorial (TUT1)

Groups	TUT1-MOV	TUT1-SEL	TUT1-MAN	TUT1-PLA	TUT1-SIN
BaseGroup vs. VideoGroup	0.9272	0.3441	0.9828	0.887	0.4275
BaseGroup vs. TextGroup	0.0287*	0.5677	0.9608	0.9992	0.9936
BaseGroup vs. VideoTextGroup	0.066	0.9899	0.7582	0.9973	0.9973
VideoGroup vs. TextGroup	0.0171*	0.9916	0.6555	0.8765	0.2511
VideoGroup vs. VideoTextGroup	0.0171*	0.3373	0.2021	0.8999	0.3692
TextGroup vs. VideoTextGroup	0.9673	0.4722	0.8506	0.9857	0.9957
Total	0.0012*	0.1857	0.3012	0.8284	0.2248

teleportation yet, so they take longer to get closer to objects. In addition, they seemed to have problems understanding the task, as they tended to select the same object repeatedly when the task specified they had to select several different objects.

For the study of the results of the third task (*TUT1-MAN*), one of the participants was considered an outlier and therefore discarded. The user's exhibited behavior was not aligned with the intended process of the experiment, as she/he did not try to complete the tasks and deviated significantly from the expected behavior. With this fix, no differences between the groups were obtained (see Table 8 and 9). This shows that the users understand the interaction with objects. It can be concluded that the previous step and this one could be merged into one as it seems that, given the context, the interactions are natural for the user. Although a new button on the controller is being used, its position on the controller seems discoverable and usable for users, allowing them to perform the new action without issues.

Selection and manipulation allow users to move forward to the placing objects in specific locations phase (*TUT1-PLA*). At this point, the user should be familiar with the main object interactions, so different ways to use them are presented. Users with a video explanation at their disposal (*VideoGroup* and *VideoTextGroup*) tend to perform this action better than the other groups (see Table 8 and 9). This suggests that if the user already knows the basis of the new interaction, they benefit more from a visual than a textual representation.

The last task of the first tutorial is about learning social interaction (*TUT1-SIN*). No difficulties with this action were observed (see Table 8 and 9). It seems that this is because interaction with NPCs is done by pressing the 'Press to talk' button. Instructions on how to interact with interfaces are provided before the experiment begins. This prior knowledge, along with the number of times the user has practiced

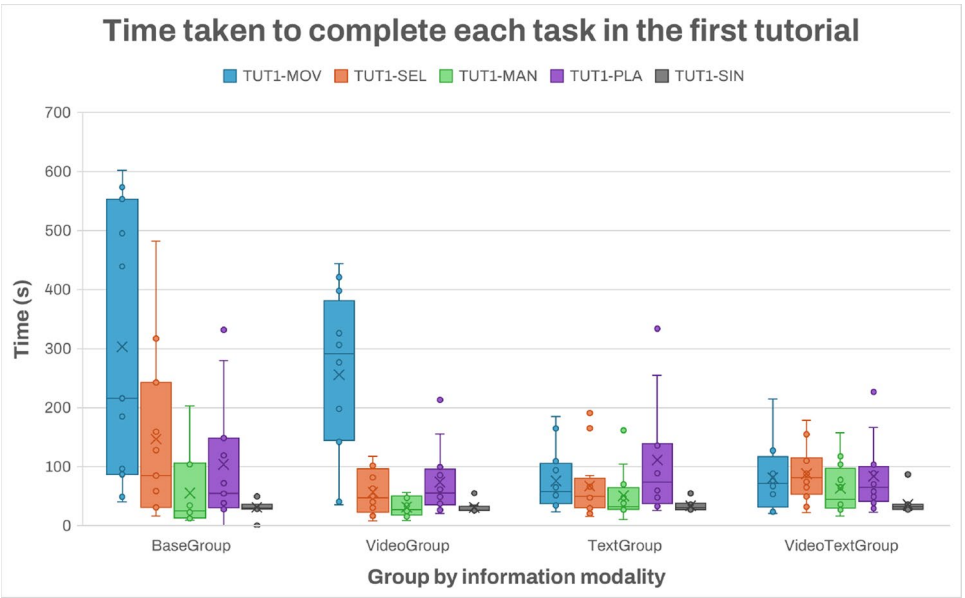

Fig. 8 Visual information showing the time distribution (in seconds) taken by users to complete each of the tasks in the first tutorial (TUT1). The users are divided into different groups based on the information modality presented to them. A large data dispersion is observed mainly in *BaseGroup* and *VideoGroup*. The most affected task is *TUT1-MOV*. The first stage of the tutorial seems to be the most problematic for non-textual explanation groups

Table 10 Results of the time taken by the user to complete the tasks proposed in the second tutorial (TUT2). These tasks focus on the interaction with the CBR vehicle (CBR), the use of the walkie-talkie for reporting (WAL), the functionality of the SABG-100 probe (SAB), the importance of the EPD dosimeter (EPD) and the request for help on demand (SOS). Values are presented as mean (standard deviation) for each group. For each factor, t-statistics, degrees of freedom, and p-values are reported

Groups	TUT2-CBR	TUT2-WAL	TUT2-SAB	TUT2-EPD	TUT2-SOS
BaseGroup	187.50 (146.56)	51.20 (11.71)	32.20 (22.99)	12.76 (2.54)	278.72 (132.20)
VideoGroup	32.69 (12.54)	44.89 (4.22)	21.61 (11.47)	10.53 (3.99)	122.12 (43.95)
TextGroup	85.67 (52.25)	59.54 (22.15)	41.70 (33.22)	17.96 (15.67)	38.05 (27.77)
VideoTextGroup	67.07 (39.02)	53.25 (16.54)	41.40 (32.64)	18.64 (12.09)	68.56 (50.52)
DF	22	22	22	22	21
t-stats	−0.946	−0.755	−0.022	0.113	1.734
p-value	0.419	0.488	1	0.525	0.097

Table 11 Kruskal–Wallis test results comparing the four groups on the time taken to complete the tasks in the second tutorial (TUT2)

Groups	TUT2-CBR	TUT2-WAL	TUT2-SAB	TUT2-EPD	TUT2-SOS
BaseGroup vs. VideoGroup	0.0013*	0.3997	0.7017	0.3367	0.9983
BaseGroup vs. TextGroup	0.5308	0.9999	0.6049	0.9347	0.0937
BaseGroup vs. VideoTextGroup	0.2846	0.9597	0.7017	0.9998	0.5611
VideoGroup vs. TextGroup	0.0337*	0.4368	0.0626	0.6847	0.002*
VideoGroup vs. VideoTextGroup	0.2511	0.9544	0.1417	0.4514	0.091
TextGroup vs. VideoTextGroup	0.8506	0.8999	1	0.9207	0.3441
Total	0.0017*	0.4141	0.052*	0.321	0.0014*

this action by activating a task or starting a video, positively influences the results. In addition, the direct action displayed on the button means that users have no problems regardless of the group to which they belong.

Users seem to have less trouble accomplishing tasks proposed in the second tutorial. Nevertheless, remarkable differences between groups can be observed (see Table 10 and Fig. 9).

The first task of the second tutorial (*TUT2-CBR*) shows the most significant difference (see Table 10). The interaction with the CBR vehicle seems to have the same problem as object placement (*TUT1-PLA*), as depicted in Table 8. This time, interaction with the interfaces is the base functionality that users need to practise. Moreover, in contrast to the social interaction phase (*TUT1-SIN*) the diversity of elements that make up this new interactive panel makes a visual explanation more necessary. Therefore, groups with visual explanations (*VideoGroup* and *VideoTextGroup*) obtain better results. Users seem to need a visual representation to show them how to interact with the CBR panel. Although the lack of video worsens the results, using text to learn about the vehicle functionality helps the results.

As shown in Table 10, users did not seem to experience significant difficulties using the walkie-talkie (action *TUT2-WAL*). Using a walkie-talkie requires knowledge about grabbing objects and interacting with interfaces, and users have already practiced these two actions several times before reaching this phase. The comparison of group results for *TUT2-WAL* does not show a significant difference.

The use of the SABG-100 probe (*TUT2-SAB*) and the EPD dosimeter (*TUT2-EPD*) is crucial, as these two tools must be used most of the time during a radioactive emergency. Therefore, the user needs to understand their purpose and how to use these devices properly. According to the time it takes for users to accomplish these tasks, their use seems reasonably straightforward (see Table 10). This suggests that the information provided by the text is not essential for completing this task. However, the number of users who used these devices in the second practice scenario (*Emergency Training*), suggests that the absence of textual explanations regarding their functionality negatively impacted their usage (see Table 12). Without a description of the device, users were unaware of its importance during a radiation emergency, leading to reduced usage. It seems that users in the text-free groups (*BaseGroup* and *VideoGroup*) tend to mistakenly believe that they should use the EPD dosimeter for this measurement when measuring the radiation rate of objects and people. Ultimately, text-free groups complete this task in less time but do not understand the purpose of these objects. This suggests that text is necessary to explain important system concepts to users, especially in systems designed to support a specific learning process.

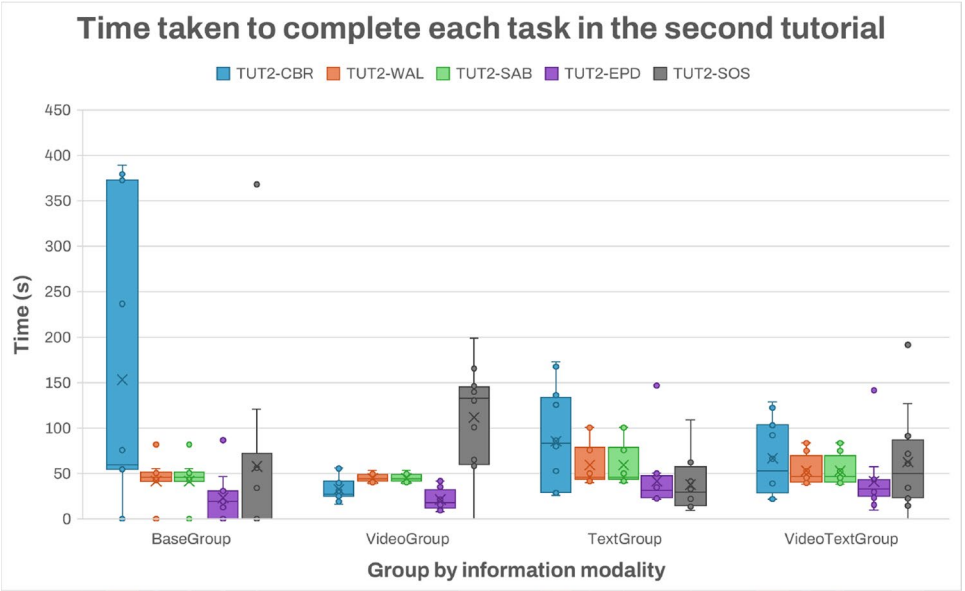

Fig. 9 Visual information showing the time distribution (in seconds) taken by users to complete each of the tasks in the second tutorial (TUT2). A large data dispersion is observed mainly in *BaseGroup*. The most affected task is *TUT2-CBR*. The first stage of the tutorial seems to be the most problematic for non-visual explanation groups

Table 12 Use of the dosimeters in the second emergency scenario, *Emergency Training*. The table shows the number of users who used these devices. It can be seen that all users provided with a textual explanation used both devices during the emergency performance. However, some users in those groups without text did not use it

Groups	No. of users (SABG-100)	No. of users (EPD)
BaseGroup	5	4
VideoGroup	8	8
TextGroup	12	12
VideoTextGroup	12	12

Table 13 Number of controller buttons pressed, *user inputs*. *Total number* shows the total number of inputs. *Movement* refers to the teleport button. *Interface* is the button used to interact with interfaces. *Manipulation* pressed to grab and release objects. *Ask Help* used to ask for help. Values are presented as mean (standard deviation) for each group. For each factor, t-statistics, degrees of freedom, and p-values are reported

Groups	Total number	Movement	Interface	Manipulation	Demand Help
BaseGroup	1023.27 (327.21)	295.91 (132.86)	473.45 (176.98)	242.09 (66.87)	15.73 (15.37)
VideoGroup	832.08 (298.35)	271.75 (79.46)	318.25 (181.16)	219.33 (131.23)	31.17 (35.32)
TextGroup	553.17 (140.34)	280.67 (61.05)	168.17 (102.70)	100.75 (52.60)	3.67 (1.70)
VideoTextGroup	654.67 (195.94)	307.08 (84.80)	225.75 (102.64)	116.75 (61.69)	6.08 (4.19)
Degrees of freedom	22	22	22	22	22
t-statistics	1.398	0.84	1.314	0.655	1.659
p-value	0.312	0.41	0.083	0.644	0.111

In the last phase, the user is explained how to request on-demand help (*TUT2-SOS*). Two extreme outliers significantly affect the average time it takes for *BaseGroup* to open the help panel. When analyzing the entire user set, the group with the worst performance is *VideoGroup* ($m = 122.12, sd = 43.95$). It seems that having only a video tends to confuse the user. *BaseGroup* users tend to experiment by pressing buttons, whereas *VideoGroup* users expect the panel to appear automatically without having to press anything.

7.1.3 User inputs

User inputs include all button presses made by the user with the virtual reality controllers. The groups without textual information are the ones that performed the highest number of button clicks overall (see Tables 13 and 14). *TextGroup* required the fewest button presses to complete the experience ($m = 553.17, sd = 140.34$), while *BaseGroup* had to press the most buttons ($m = 1023, sd = 327.21$). *BaseGroup* also has the lowest average number of completed objectives. As a result, its ratio of button presses per completed objective is significantly higher than that of the other groups.

The results on interaction with objects and movement show how many user inputs had no real effect. To assess this, the number of button presses that resulted in an actual action (i.e. they had an actual effect in the scenario) was compared to the number of times the user picked up an object or moved by teleport. These unnecessary inputs correspond to button presses that did not trigger an action (see Fig. 10). For movement, in most cases, users pressed the button and successfully moved. The instances where users failed to

Table 14 Kruskal–Wallis test results comparing the four groups on the number of controller button presses, *user inputs*

Groups	Total number	Movement	Interface	Manipulation	Demand Help
BaseGroup vs. VideoGroup	0.5284	0.9926	0.1555	0.5478	0.9602
BaseGroup vs. TextGroup	0.0092*	0.9948	0.002*	0.004*	0.0884
BaseGroup vs. VideoTextGroup	0.066	0.9993	0.0092*	0.0067*	0.3082
VideoGroup vs. TextGroup	0.0337*	0.9916	0.0673	0.0286*	0.0463*
VideoGroup vs. VideoTextGroup	0.4368	0.93	0.6555	0.1601	0.3622
TextGroup vs. VideoTextGroup	0.7432	0.9206	0.3066	0.9673	0.4928
Total	0.0025*	0.9217	0.0003*	0.0004*	0.0184*

move happened when they attempted to teleport outside the scene boundaries. However, the differences become more evident when analyzing inputs related to object interaction. Based on these data, it can be concluded that users without textual information press more buttons randomly, as they are unaware of their actual functionality.

Groups with textual information use nearly half of their inputs for movement (see Fig. 11). However, in the other groups, the highest percentage of inputs corresponds to the button used for interacting with user interfaces. Observations during the experimental phase suggest that this happens because, at the beginning of the experiment, users tend to press this button exclusively, as it is the only one they are familiar with. As mentioned earlier, interaction with the interface is explained outside the experience, so users know how to activate videos and tasks. This leads users to overlook the possibility of other buttons on the controller, something that the textual explanation clarifies. This confusion causes users to repeatedly press the only button they know until they eventually discover the existence of other buttons, such as the one needed for teleportation.

One notable finding is that video explanations seem to encourage a more balanced use of both the left and right hands. The *BaseGroup* used the right hand 57.8% of the time, the *VideoGroup* 53.56%, the *TextGroup* 77.99%, and the *VideoTextGroup* 65.52%. This might be because the text does not specify that both hands can be used for actions, whereas the video visually shows their use. Text appears to have an influence, as *VideoTextGroup* uses the right hand more frequently than *VideoGroup*. Interestingly, *BaseGroup* shows results very similar to *VideoGroup*. The data and user behavior observations suggest this happens because users press buttons on both hands randomly. This conclusion is supported by the proportion of right and left-hand usage in object interactions, where *BaseGroup* tends to favor the right hand (see section 7.1.5). This states that users who do not receive explanations press buttons randomly with either hand but prioritize the right hand for effective actions.

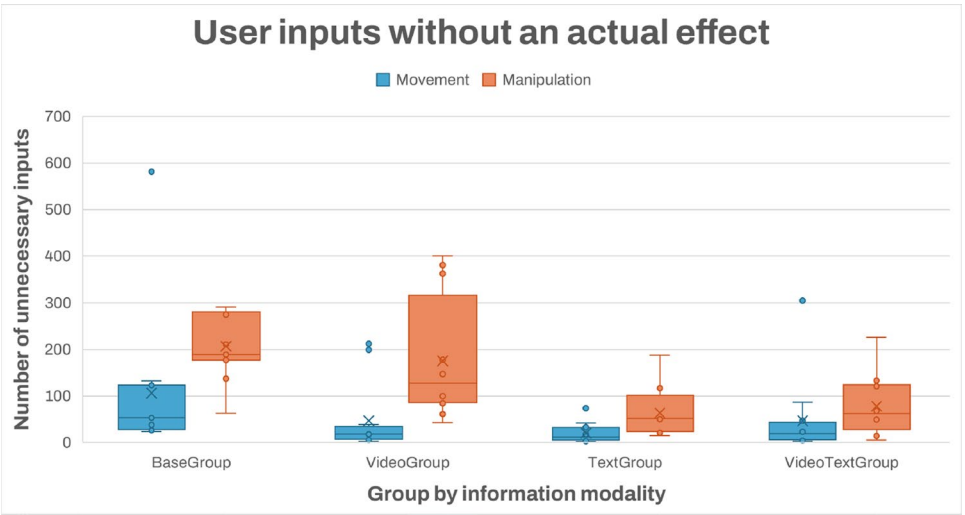

Fig. 10 Buttons that were pressed unnecessarily. *Movement* shows those failed teleport attempts. *Manipulation* the number of failed attempts to grab an object

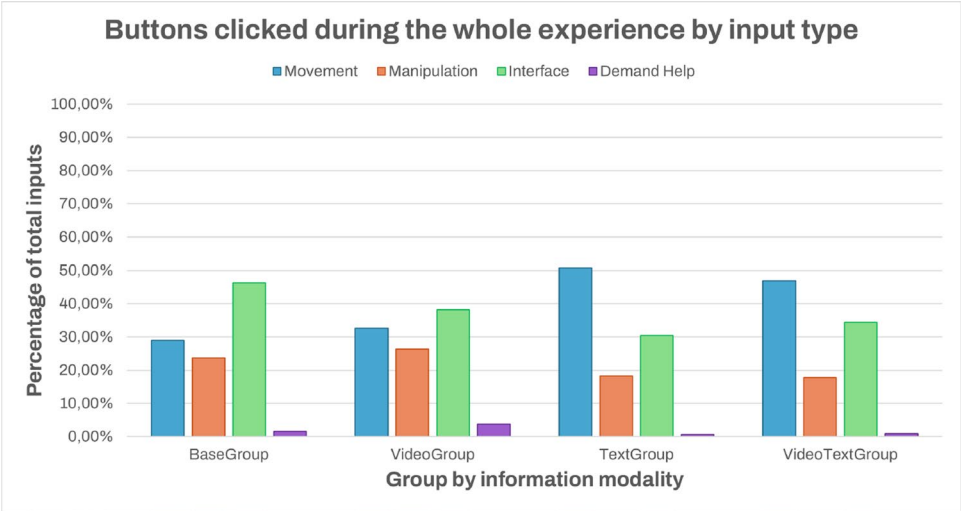

Fig. 11 Percentage of user inputs on total number of inputs. *Movement* is related to the button used to perform the virtual movement. *Interface* is the button to be pressed by the user to interact with the interface. *Manipulation* is related to the button that must be held down in order to grab an object. *Ask Help* is used to ask for help

Table 15 Displacement performed by the user during the virtual experience, both in the real world and virtual reality. This table shows the number of displacements performed and the distance. Values are presented as mean (standard deviation) for each group. For each factor, t-statistics, degrees of freedom, and p-values are reported

	No. of displacements		Total distance	
Groups	Real life	Virtual reality	Real life	Virtual reality
BaseGroup	173.83 (120.40)	876.30 (451.49)	512.20 (360.92)	151.75 (78.54)
VideoGroup	224.83 (81.74)	758.30 (237.49)	715.00 (286.18)	121.36 (42.31)
TextGroup	260.42 (49.95)	669.30 (107.26)	784.60 (122.15)	110.05 (16.06)
VideoTextGroup	260.08 (78.89)	830.60 (275.93)	769.70 (288.63)	140.58 (46.77)
Degrees of freedom	22	22	22	22
t-statistics	2.11	1.57	2.2	2.35
p-value	0.07	0.93	0.04	0.87

Table 16 Kruskal–Wallis test results comparing the four groups on the number of displacements and total distance

	No. of displacements		Total distance	
Groups	Real life	Virtual reality	Real life	Virtual reality
BaseGroup vs. VideoGroup	0.1764	0.8397	0.1033	0.6071
BaseGroup vs. TextGroup	0.0112*	0.5677	0.0092*	0.4143
BaseGroup vs. VideoTextGroup	0.3785	0.5284	0.3785	0.4143
VideoGroup vs. TextGroup	0.7258	0.6555	0.9389	0.8765
VideoGroup vs. VideoTextGroup	0.9544	0.6734	0.7921	0.9673
TextGroupvs. VideoTextGroup	0.5084	0.9998	0.3373	0.9994
Total	0.0206*	0.3402	0.0125*	0.3201

7.1.4 Movement

Studying the influence of information modality on movement is crucial to understanding how users best comprehend the space they are in. *Radiation Room* is designed for movement via teleportation. Additionally, users can move in the real world, though with certain spatial limitations. For this reason, data on their movement in both the virtual environment and the real world have been collected (see Tables 15 and 16). The data reveal that the most significant difference emerges when comparing the *BaseGroup* to the other groups. *BaseGroup* showed more movement within virtual reality and less in the real world, both in terms of the number of movements and the distance displaced. These results might suggest that this group correctly understood the teleportation functionality and, as a result, used it more efficiently to reduce physical displacement. However, a qualitative analysis of their behavior during the experiment reveals a different explanation. The lack of guidance led participants in this group to remain mostly stationary in the real world due to uncertainty about their objectives. Additionally, this same lack of information caused an erratic movement through the virtual environment. The other groups showed results that did not appear to be directly related to the type of instructional modality they received. Observations during the experiment indicate that the primary challenge for all users was the time it took to complete the tutorial section dedicated to explaining movement (see section 5.1). Once participants completed that task, movement no longer posed a difficulty. Therefore, we can conclude that user movement is positively influenced by the presence of textual explanations, even if this influence is not apparent later on, and that the absence of information during the virtual experience negatively affects user behavior.

Table 17 Number of interactions with objects. *Total number* shows the total number interactions. The rest of the columns show a division between the two types of interactions the user can perform (*selection* and *manipulation*). Values are presented as mean (standard deviation) for each group. For each factor, t-statistics, degrees of freedom, and p-values are reported

Groups	No. of interactions	Selection	Manipulation
BaseGroup	213.18 (94.68)	71.73 (35.60)	35.09 (14.09)
VideoGroup	252.42 (71.68)	82.83 (25.52)	43.67 (11.85)
TextGroup	232.83 (42.11)	79.17 (20.47)	37.33 (4.09)
VideoTextGroup	276.33 (101.14)	99.67 (41.34)	38.58 (12.47)
Degrees of freedom	22	22	22
t-statistics	1.32	1.47	0.32
p-value	0.73	0.53	0.39

Table 18 *Kruskal–Wallis* test results comparing the four groups on the number of interaction with objects

Groups	No. of interactions	Selection	Manipulation
BaseGroup vs. VideoGroup	0.722	0.8542	0.6649
BaseGroup vs. TextGroup	0.9167	0.6844	1
BaseGroup vs. VideoTextGroup	0.758	0.3957	0.9999
VideoGroup vs. TextGroup	0.982	0.9998	0.7589
VideoGroup vs. VideoTextGroup	0.9939	0.9105	0.5604
TextGroup vs. VideoTextGroup	0.9857	0.9207	0.8204
Total	0.7082	0.4579	0.5128

7.1.5 Item interactions

The number of interactions performed by users suggests that there is no significant difference in object interaction (see Tables 17 and 18). Overall, users tend to be efficient in their usage, interacting with objects only when it is necessary. The lower number of interactions in *BaseGroup* may be related to the number of objectives they have completed.

As mentioned earlier, significant differences can be observed in the use of the left and right hands (see Table 19). Groups that received an explanation, whether through video or text, tend to use both hands interchangeably. In particular, those who can view the explanation through a video show a greater similarity, possibly driven by the fact that the videos demonstrate interactions using both hands.

When analyzing how users use their hands, there seems to be a difference in the groups in which users had text. These groups tend to use the left hand more often for detecting objects than for grabbing them. This significant difference does not occur with *VideoGroup*.

As observed, *BaseGroup* has the largest difference in the percentage of use between both hands. Previous analysis of user inputs shows that this does not happen with button presses (see section 7.1.3). Analyzing this difference leads to the conclusion that *BaseGroup* users press buttons randomly, but once they discover the functionality, they prefer to use their right hand to perform an effective action.

7.2 Usability, user experience and cognitive load results

The results of the user experience and usability evaluation, conducted using the UEQ, show no significant difference between the modalities (see Table 20). However, the *perspicuity* factor, which focuses on familiarity and learning of the tool, shows a difference between the groups with text and those without. This states users rate text-based explanations positively.

Table 19 Percentages of left and right hand use to interact with objects. The first two columns show information about all interactions. The *selection* columns show the information related to the detection of interactable objects. The *manipulation* columns are related to the interactions performed to grab objects. The use of both hands indistinctly is observed in those groups where interactions are explained by videos

	Total		Selection		Manipulation	
Groups	Left	Right	Left	Right	Left	Right
BaseGroup	26.10%	73.90%	26.62%	73.38%	24.87%	75.13%
VideoGroup	48.26%	51.73%	50.40%	49.60%	44.47%	55.53%
TextGroup	30.42%	69.58%	35.26%	64.73%	20.09%	79.91%
VideoTextGroup	40.17%	59.83%	42.89%	57.11%	33.26%	66.74%

Table 20 User experience results based on the factors proposed by the *User Experience Questionnaire*. Values are presented as mean (standard deviation) for each group. For each factor, t-statistics, degrees of freedom, and p-values are reported

Groups	Attractiveness	Perspicuity	Efficiency	Dependability	Stimulation	Novelty
BaseGroup	0.15 (0.87)	−0.16 (0.47)	0.16 (0.93)	−0.25 (1.04)	0.30 (1.00)	0.39 (0.62)
VideoGroup	−0.24 (0.78)	−0.52 (0.47)	−0.08 (0.60)	−0.60 (0.74)	0.13 (0.93)	0.35 (0.56)
TextGroup	−0.13 (0.63)	0.27 (0.37)	0.10 (0.75)	−0.17 (0.73)	−0.08 (0.69)	0.25 (0.37)
VideoTextGroup	−0.12 (0.80)	0.16 (0.60)	0.11 (0.70)	−0.34 (0.89)	0.18 (0.98)	0.20 (1.04)

Table 21 Descriptive statistics of NASA-TLX traits for each experimental group according to information modality. Values are presented as mean (standard deviation) for each group. For each factor, t-statistics, degrees of freedom, and p-values are reported

Trait	BaseGroup	VideoGroup	TextGroup	Video-TextGroup
Mental Demand	75.83 (15.64)	56.67 (21.03)	58.33 (15.28)	57.27 (24.94)
Physical Demand	24.17 (13.79)	20.83 (10.84)	28.33 (25.17)	30.91 (19.73)
Temporal Demand	40.00 (22.96)	25.83 (21.52)	45.00 (22.76)	47.27 (19.02)
Effort	72.5 (21.79)	47.5 (22.21)	45.0 (28.12)	45.5 (19.16)
Performance	55.00 (27.47)	71.67 (25.17)	83.33 (9.85)	73.64 (23.78)
Frustration	59.67 (10.12)	51.83 (12.87)	60.67 (14.10)	55.64 (15.26)

The results of the cognitive load evaluation, conducted using the NASA-TLX, are shown in Table 21 and Fig. 12. The final cognitive load for the different groups shows that users tend to have very similar loads regardless of the group. However, certain factors show some differences between them. *BaseGroup* is the group for which the virtual reality experience has resulted in the highest mental load ($m = 75.83$). The lack of knowledge about the actions they need to perform seems to lead to a higher mental load as they have to figure out everything related to the environment. This group perceives having had the weakest performance in completing the set objectives ($m = 55$). In this group, certain users consider their performance to be very negative, something not observed in the other groups. The results of the other groups tend to be better, indicating that the overall feeling of these users is generally positive. *BaseGroup* alose reported experiencing very high effort ($m = 72.5$). Similar to mental demand, this could be due to the constant need to figure out what they are supposed to do. The frustration factor, on the other hand, did not show significant results. At the beginning of the experiment, the users were reminded to stay calm, that they simply had to do what they thought was right without feeling judged. This may have led the user not to feel frustration, thinking that it doesn't matter if they are doing it right or wrong, as it is part of the experience. Another noteworthy fact is that *VideoGroup* considers that the *temporal demand* is low ($m = 25.83$). It seems that the exclusive use of video makes the user perceive the experience as more efficient. However, as observed when analyzing the use of dosimeters, temporal efficiency does not imply better learning.

8 Discussion

This study analyzed the impact of using text and video as instructional material for a learning process in a virtual reality environment. Four different modalities have been studied: no information, text, video, and both text and video. The results show that the absence of information negatively impacts user performance. The lack of explanations also slightly affects user experience and cognitive load. In addition, a significant difference in the impact of the information modality could be observed about the explained functionality, which supports the main hypothesis of this work.

These findings are consistent with previous studies that have pointed out the importance of initial orientation in virtual reality environments (Miguel-Alonso et al. 2024; Villar and León 2023a). Previous research has found that in

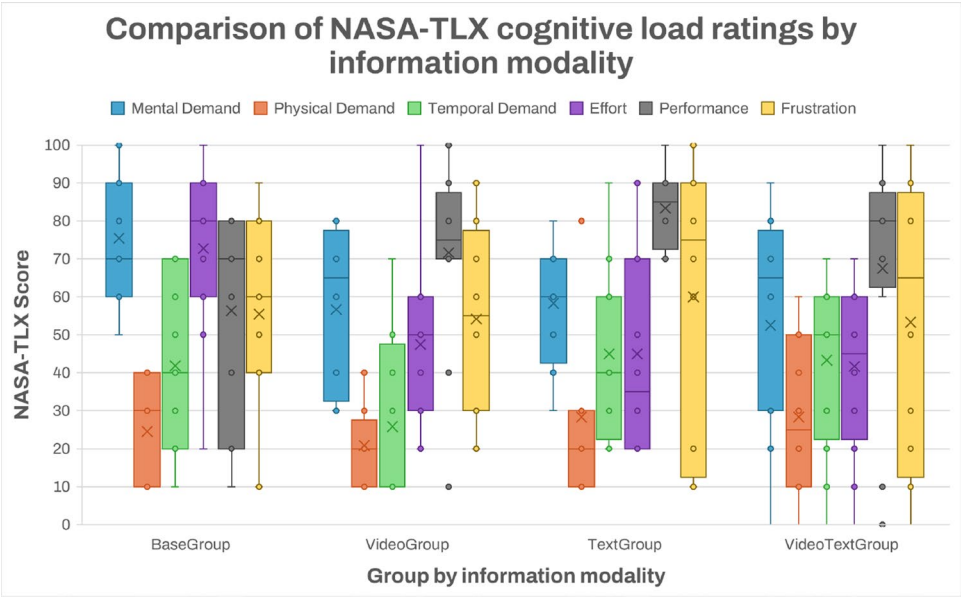

Fig. 12 NASA-TLX workload scores across experimental groups

most cases, information is presented in textual format (Kao et al. 2021), but visual demonstrations can also have a positive impact (Thoravi Kumaravel et al. 2019). However, in contrast to expectations, combining both modalities did not show a significant advantage over using only one, suggesting that redundancy in information presentation may not generate additional benefits (Mayer 2014).

8.1 Influence of tutorial modality on user actions in virtual environments

The results show that no modality consistently yields more positive results than the others. Previous research has shown that interactive tutorials have a significantly greater impact on learning outcomes compared to non-interactive formats (Kao 2020). The presented tutorial was designed to be fully interactive from the outset. Within this interactive context, we examined the influence of information modality (text vs. video). The limited differences observed between the two modalities can be explained by the fact that the interactivity itself drives user learning, making the specific format of information modality (text or video) less critical. During the analysis, it was observed that the positive effect of a modality depends on the type of explanation provided during tutoring. Text seems to be more effective for explaining new and specific concepts. Positive results were also observed when the functionality of a specific object is important, and the user needs to understand it in order to use it, such as in the case of dosimeters. On the other hand, video is more effective for the learning of actions that visually demonstrate a variation of an action already explained. For example, placing objects in specific locations only requires knowing the visual representation of the place where the object should be placed. Neither video nor text always achieves the best results, indicating that learning is more beneficial with a single modality for each particular category of information. Combining video and text for a single information unit may lead to worse outcomes than using the most effective modality for that particular explanation. In general, the group with only text (*TextGroup*) shows better results than video (*VideoGroup*), although the statistical results did not achieve numerical significance. The group that did not get explanations in the tutorial (*BaseGroup*) obtains the worst results. As Kao noted, using some support is always better than nothing, regardless of the modality used (Kao 2020). Another conclusion to highlight is that the results show a greater difference during the technology adaptation phase, with more similar results during the acquisition of basic emergency concepts.

Interaction with the elements shows that, regardless of the modality, users efficiently use the objects and only interact with them when necessary. However, the presence of video leads to an increase in the use of the left hand. This is because the video allows the user to visualize the use of both hands interchangeably. An analysis of the number of controller buttons pressed shows that the group with text (*TextGroup*) required the fewest presses. This efficiency is based on how many presses the user made to conduct a real action, such as using teleportation or grabbing objects. Groups without text were more prone to make mistakes in pressing the correct buttons to perform actions. It is assumed that users press buttons randomly until they find the ones that allow them to perform the action they want at that moment. Furthermore, the lack of textual information causes these users to press the button to interact with the interface more often, as it is the only button they know at the beginning of the scenarios. The modality slightly affects movement, with users who have textual information making more use of teleportation. It is important to take this into account in cases where the user may not have a spacious area available during their learning experience.

8.2 User perception based on the modality

User experience does not show significant differences regarding modality. However, the *perspicuity* factor, related to familiarization and learning how to use the tool, yields more positive results for those groups with textual explanations only. Regarding cognitive load, the absence of explanations tends to affect mental load negatively, performance, and effort. Additionally, user behavior observations show signs of frustration during the experiment. Temporal demand positively affects the group with video (*VideoGroup*) perception. However, while using videos requires less temporal effort, it negatively impacts the knowledge acquired, as seen in the use of dosimeters. These results show that the modality does not have a significant impact, but the lack of tutoring does have a negative influence (Lee et al. 2024).

8.3 Importance of information modality in a tutoring process and limitations

These findings can inform the design of more effective learning processes in a virtual reality environment. In particular, the results suggest that the tutoring process is almost indispensable (Kao 2020), and it must consider what concept wants to be explained to choose the appropriate information modality.

This study has some limitations that should be considered when interpreting the results. First, the virtual reality scenarios used in this study are complex. The large number of actions that can be performed in it may affect the dispersion of the data obtained from the users. The results related to movement may have been affected by the space

limitations of the real world and the wired virtual reality system, where users had very limited freedom of movement (Shewaga et al. 2020). More space and a wireless set could allow for better observation of how the user moves independently. Second, although the participants in the study had no prior relationship with virtual reality, their profiles may have affected the final results (Petri et al. 2020). For example, people who are used to general technology may perform better than those who have little or no connection to the technological world. Finally, the qualitative observations of the videos suggest that real user behavior does not match their results in the questionnaires related to user experience and cognitive load. During the experiment, it could be observed by the researchers that some users were frustrated by not being able to do certain actions. Further research is required to confirm or reject this observation.

9 Conclusions and future work

Virtual reality learning environments can help users to acquire basic concepts efficiently, especially in scenarios in which the alternatives are costly or dangerous. To use virtual reality technology correctly and focus on the learning objectives, users must be familiar with the interaction methods, which can be achieved through a contextualized interactive tutoring phase. However, the explanations shown in these phases can be presented in different modalities, and this can affect their effectiveness. In this study, we have analyzed the impact of a tutoring phase in which different groups received the information with different modalities, and compared the relative impact of the use of text and video when explaining different categories of information. The results show that the modality of the information affects it in different ways, depending on the explanation. Text is more effective for new functionalities not previously practiced and new concepts, and video is more effective in explanations that present variations of functionalities already explained. Combining both modalities did not result in a significant advantage, supporting the idea that redundancy is not necessary. User experience and cognitive load are also slightly affected by these variations. The results evidence that this tendency is constant regardless of the previous knowledge of the learning objectives. These conclusions strongly support the initial hypotheses, namely that text is better for descriptive concepts and video is better for explanative in virtual reality tutorials.

Based on these conclusions, further aspects can be explored. First, it is necessary to observe the persistence of learning over time. To do so, repeating the a similar set of scenarios with same subjects can allow for an analysis of their memory and learning. Furthermore, it could be beneficial to gain more insight on the impact of cognitive load on the learning process. The behavior qualitatively observed during the experimental phase differs from the results obtained in the questionnaires. Second, it is important to more precisely analyze impact that the user's profile has on the learning process. Familiarity with the used technology could greatly affect the use of virtual reality, regardless of the learning modality. Finally, the specific content of the videos and text fragments could benefit from a specific study for comparing which messages or constructions are best. For instance, comparing whether showing the information in the videos in first or third person is best, or comparing the effect of addressing the user by her/his name in the text instead of generically.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10055-025-01240-y.

Author contributions All authors contributed equally to this work.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no Conflict of interest.

References

Andersen E, O'Rourke E, Liu, Y-E, Snider R, Lowdermilk J, Truong D, Cooper S, Popovic Z (2012) The impact of tutorials on games of varying complexity. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '12, pp. 59–68. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/2207676.2207687 . https://dl.acm.org/doi/10.1145/2207676.2207687 Accessed 27 June 2023

Antonenko P, Paas F, Grabner R, Van Gog T (2010) Using electroencephalography to measure cognitive load. Educ Psychol Rev 22:425–438

Albus P, Vogt A, Seufert T (2021) Signaling in virtual reality influences learning outcome and cognitive load. Comput Educ 166:104154

Bhargava A, Bertrand JW, Gramopadhye AK, Madathil KC, Babu SV (2018) Evaluating multiple levels of an interaction fidelity continuum on performance and learning in near-field training simulations. IEEE Trans Vis Comput Graph 24(4):1418–1427

Boletsis C, Chasanidou D (2022) A typology of virtual reality locomotion techniques. Multimodal Technol Interact 6(9):72

Bevana N, Kirakowskib J, Maissela J (1991) What is usability. In: Proceedings of the 4th International Conference On HCI, Stuttgart, Germany, pp. 1–6. https://usabilitynet.org/papers/whatis92.pdf

Bozgeyikli E, Raij A, Katkoori S, Dubey R (2016) Point & Teleport Locomotion Technique for Virtual Reality. In: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play. CHI PLAY '16, pp. 205–216. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/2967934.2968105 . https://dl.acm.org/doi/10.1145/2967934.2968105 Accessed 03 July 2023

Brooke J (1995) SUS: a quick and dirty usability scale. Usability Eval Ind 189:4–7

Checa D, Bustillo A (2020) A review of immersive virtual reality serious games to enhance learning and training. Multimed Tools Appl 79(9):5501–5527

Clifford RMS, Jung S, Hoermann S, Billinghurst M, Lindeman RW (2019) Creating a Stressful Decision Making Environment for Aerial Firefighter Training in Virtual Reality. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 181–189. https://doi.org/10.1109/VR.2019.8797889 . ISSN: 2642-5254

Checa D, Miguel-Alonso I, Bustillo A (2021) Immersive virtual-reality computer-assembly serious game to enhance autonomous learning. Virtual Reality. https://doi.org/10.1007/s10055-021-00607-1. Accessed 26 June 2023

Cherni H, Métayer N, Souliman N (2020) Literature review of locomotion techniques in virtual reality. Int J Virtual Real EBSCOhost. https://doi.org/10.20870/IJVR.2020.20.1.3183

Damen D, Leelasawassuk T, Haines O, Calway A, Mayol-Cuevas WW (2014) You-Do, I-Learn: discovering task relevant objects and their modes of interaction from multi-user egocentric video. In: BMVC, vol. 2, p. 3

Dixon WJ, Massey Jr, FJ (1951) Introduction to Statistical Analysis. Introduction to statistical analysis. McGraw-Hill, New York, NY, US. Pages: x, 370

Dillman KR, Mok TTH, Tang A, Oehlberg L, Mitchell A (2018) A Visual Interaction Cue Framework from Video Game Environments for Augmented Reality. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18, pp. 1–12. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3173574.3173714 . https://doi.org/10.1145/3173574.3173714 Accessed 15 Sep 2023

Engelbrecht H, Lindeman RW, Hoermann S (2019) A SWOT Analysis of the Field of Virtual Reality for Firefighter Training. Frontiers in Robotics and AI 6. Accessed 16 Sep 2023

Epic Games: Unreal Engine. https://www.unrealengine.com. Version 5 (2025)

Fowler A, Cusack B (2014) A proposed method for measuring learning in video games. GSTF J Comput (JoC) 3(4):39

Frommel J, Fahlbusch K, Brich J, Weber M (2017) The Effects of Context-Sensitive Tutorials in Virtual Reality Games. In: Proceedings of the Annual Symposium on Computer-Human Interaction in Play, pp. 367–375. ACM, Amsterdam The Netherlands. https://doi.org/10.1145/3116595.3116610 . https://dl.acm.org/doi/10.1145/3116595.3116610 Accessed 2023-06-26

Farra SL, Gneuhs M, Hodgson E, Kawosa B, Miller ET, Simon A, Timm N, Hausfeld J (2019) Comparative cost of virtual reality training and live exercises for training hospital workers for evacuation. CIN Comput Inform Nurs 37(9):446–454

Gordillo A, Barra E, Aguirre S, Quemada J (2014) The usefulness of usability and user experience evaluation methods on an e-learning platform development from a developer's perspective: A case study. In: 2014 IEEE Frontiers in Education Conference (FIE) Proceedings, pp. 1–8. IEEE

Gee JP (2005) Learning by design: good video games as learning machines. E-Learn Digit Med 2(1):5–16

Gavish N, Gutiérrez T, Webel S, Rodríguez J, Peveri M, Bockholt U, Tecchia F (2015) Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. Interact Learn Environ 23(6):778–798

Green M, Khalifa A, Barros G, Togellius J (2017) "press space to fire": automatic video game tutorial generation. Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment 13(2):75–80. https://doi.org/10.1609/aiide.v13i2.12977. Number: 2. Accessed 22 Sep 2023

Gürer S, Surer E, Erkayaoğlu M (2023) MINING-VIRTUAL: A comprehensive virtual reality-based serious game for occupational health and safety training in underground mines. Safety Science 166:106226. https://doi.org/10.1016/j.ssci.2023.106226. Accessed 26 July 2023

Goto M, Uematsu Y, Saito H, Senda S, Iketani A (2010) Task support system by displaying instructional video onto AR workspace. In: 2010 IEEE International Symposium on Mixed and Augmented Reality, pp. 83–90. https://doi.org/10.1109/ISMAR.2010.5643554

Ghosh S, Winston L, Panchal N, Kimura-Thollander P, Hotnog J, Cheong D, Reyes G, Abowd GD (2018) NotifiVR: Exploring Interruptions and Notifications in Virtual Reality. IEEE Trans. Vis. Comput. Graph. 24(4):1447–1456. https://doi.org/10.1109/TVCG.2018.2793698. Conference Name: IEEE Transactions on Visualization and Computer Graphics. Accessed 02 Oct 2023

Hart SG (2006) Nasa-task load index (nasa-tlx); 20 years later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 50(9):904–908. https://doi.org/10.1177/154193120605000909. Publisher: SAGE Publications Inc., Accessed 14 Nov 2023

Ho JCF (2017) Practice in Reality for Virtual Reality Games: Making Players Familiar and Confident with a Game. In: Bernhaupt, R., Dalvi, G., Joshi, A., K. Balkrishan, D., O'Neill, J., Winckler, M. (eds.) Human-Computer Interaction - INTERACT 2017. Lecture Notes in Computer Science, pp. 147–162. Springer, Cham. https://doi.org/10.1007/978-3-319-67684-5_10

Hassenzahl M, Tractinsky N (2006) User experience-a research agenda. Behaviour Inform Technol 25(2):91–97

IAEA IAEA (2005) Radiological Emergency Response Exercise. https://www-pub.iaea.org/mtcd/publications/pdf/eprexert/PDF/MedExManual.pdf Accessed 26 March 2025

Isaksson M, Raaf CL (2017) Environmental Radioactivity and Emergency Preparedness. CRC Press, Boca Raton.https://doi.org/10.1201/9781315372877

Jensen L, Konradsen F (2018) A review of the use of virtual reality head-mounted displays in education and training. Educ Inf Technol 23:1515–1529

Kao D (2020) Exploring Help Facilities in Game-Making Software. In: Proceedings of the 15th International Conference on the Foundations of Digital Games. FDG '20, pp. 1–14. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3402942.3403014 . https://dl.acm.org/doi/10.1145/3402942.3403014 Accessed 2025-07-22

Kirschner PA (2002) Cognitive load theory: Implications of cognitive load theory on the design of learning. Elsevier, Amsterdam

Kao D, Magana AJ, Mousas C (2021) Evaluating Tutorial-Based Instructions for Controllers in Virtual Reality Games. Proceedings of the ACM on Human-Computer Interaction 5(CHI PLAY),

234–123428 https://doi.org/10.1145/3474661 . Accessed 26 June 2023
Kleven NF, Prasolova-Førland E, Fominykh M, Hansen A, Rasmussen G, Sagberg LM, Lindseth F (2014) Training nurses and educating the public using a virtual operating room with Oculus Rift. In: 2014 International Conference on Virtual Systems & Multimedia (VSMM), pp. 206–213. https://doi.org/10.1109/VSMM.2014.7136687
LaValle SM (2023) Virtual Real. Cambridge University Press, Cambridge
Lele A (2013) Virtual reality and its military utility. J Ambient Intell Humaniz Comput 4(1):17–26
Laugwitz B, Held T, Schrepp M (2008) Construction and Evaluation of a User Experience Questionnaire. In: Holzinger, A. (ed.) HCI and Usability for Education And Work, pp. 63–76. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-89350-9_6
Lee Y, Kim G, Lee KH, Park J, Kim HK (2024) Comparison of tutorial methods in virtual reality games for a better user experience. Appl Sci 14(16):7141. https://doi.org/10.3390/app14167141. Number: 16 Publisher: Multidisciplinary Digital Publishing Institute. Accessed 04 March 2025
Lackey SJ, Salcedo J, Szalma JL, Hancock PA (2016) The stress and workload of virtual reality training: the effects of presence, immersion and flow. Ergonomics 59(8):1060–1072
Miguel-Alonso I, Checa D, Guillen-Sanz H, Bustillo A (2024) Evaluation of the novelty effect in immersive virtual reality learning experiences. Virtual Real 28(1):27
Miguel-Alonso I, Rodriguez-Garcia B, Checa D, Bustillo A (2023) Countering the novelty effect: a tutorial for immersive virtual reality learning environments. Appl Sci 13(1):593
Mayer RE (1995) Structural analysis of science prose: can we increase problem-solving performance? ACM SIGDOC Ast J Comput Doc 19(3):3–25
Mayer RE (2014) Cognitive Theory of Multimedia Learning. In: Mayer, R.E. (ed.) The Cambridge Handbook of Multimedia Learning, 2nd edn. Cambridge Handbooks in Psychology, pp. 43–71. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781139547369.005.https://www.cambridge.org/core/books/cambridge-handbook-of-multimedia-learning/cognitive-theory-of-multimedia-learning/24E5AEDEC8F4137E37E15BD2BCA91326 Accessed 2023-06-27
Meta: First Steps on Meta Quest — meta.com. https://www.meta.com/es-es/experiences/1863547050392688/. [Accessed 04-03-2025] (2019)
Makransky G, Petersen GB (2021) The cognitive affective model of immersive learning (Camil): a theoretical research-based model of learning in immersive virtual reality. Educ Psychol Rev 33(3):937–958
Makransky G, Terkildsen TS, Mayer RE (2019) Adding immersive virtual reality to a science lab simulation causes more presence but less learning. Learn Instr 60:225–236
Nassar AK, Al-Manaseer F, Knowlton LM, Tuma F (2021) Virtual reality (VR) as a simulation modality for technical skills acquisition. Ann Med Surg. 71:102945
Narciso D, Melo M, Raposo JV, Cunha J, Bessa M (2020) Virtual reality in training: an experimental study with firefighters. Multimed Tools Appl 79(9):6227–6245
Naranjo JE, Sanchez DG, Robalino-Lopez A, Robalino-Lopez P, Alarcon-Ortiz A, Garcia MV (2020) A scoping review on virtual reality-based industrial training. Appl Sci 10(22):8224
Petri K, Feuerstein K, Folster S, Bariszlovich F, Witte K (2020) Effects of Age, Gender, Familiarity with the Content, and Exposure Time on Cybersickness in Immersive Head-mounted Display Based Virtual Reality
Pellas N, Fotaris P, Kazanidis I, Wells D (2019) Augmenting the learning experience in primary and secondary school education: a systematic review of recent trends in augmented reality game-based learning. Virtual Real 23(4):329–346
Petersen N, Stricker D (2012) Learning task structure from video examples for workflow tracking and authoring. In: 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 237–246. https://doi.org/10.1109/ISMAR.2012.6402562
Rettinger M, Hug M, Kamel H, Saxena Y, Rigoll G (2023) Enhancing VR Training: Impact of Information Transfer Methods. In: 2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), pp. 513–518. https://doi.org/10.1109/ISMAR-Adjunct60411.2023.00111 . ISSN: 2771-1110
Radhakrishnan U, Koumaditis K, Chinello F (2021) A systematic review of immersive virtual reality for industrial skills training. Behav Inform Technol 40(12):1310–1339
Radianti J, Majchrzak TA, Fromm J, Wohlgenannt I (2020) A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. Comput Educ 147:103778
Rodrigues L, Pereira FD, Toda AM, Palomino PT, Pessoa M, Carvalho LSG, Fernandes D, Oliveira EHT, Cristea AI, Isotani S (2022) Gamification suffers from the novelty effect but benefits from the familiarization effect: findings from a longitudinal study. Int J Educ Technol High Educ 19(1):13
Sun W, Chen Q (2023) The Design, Implementation and Evaluation of Gamified Immersive Virtual Reality (IVR) for Learning: A Review of Empirical Studies. Euro Conference on Games Based Learn 17(1):789–797
Somrak A, Humar I, Hossain MS, Alhamid MF, Hossain MA, Guna J (2019) Estimating VR Sickness and user experience using different HMD technologies: an evaluation study. Future Gener Comput Syst 94:302–316
Sakowitz SM, Inglehart MR, Ramaswamy V, Edwards S, Shoukri B, Sachs S, Kim-Berman H (2020) A comparison of two-dimensional prediction tracing and a virtual reality patient methods for diagnosis and treatment planning of orthognathic cases in dental students: a randomized preliminary study. Virtual Real 24(3):399–409
Shewaga R, Uribe-Quevedo A, Kapralos B, Alam F (2020) A comparison of seated and room-scale virtual reality in a serious game for epidural preparation. IEEE Trans Emerg Top Comput 8(1):218–232. https://doi.org/10.1109/TETC.2017.2746085. Conference Name: IEEE Transactions on Emerging Topics in Computing
Singh B, Vig K, Kaunert C (2024) Modernizing healthcare: application of augmented reality and virtual reality in clinical practice and medical education. Modern Technology in Healthcare and Medical Education: Blockchain. IoT, AR, and VR. IGI Global Scientific Publishing, USA, pp 1–21
Sweller J (2011) Cognitive load theory. In: Psychology of Learning and Motivation vol. 55, pp. 37–76. Elsevier, The Netherlands
Thoravi Kumaravel B, Nguyen C, DiVerdi S, Hartmann B (2019) TutoriVR: A Video-Based Tutorial System for Design Applications in Virtual Reality. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. CHI '19, pp. 1–12. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3290605.3300514 . https://dl.acm.org/doi/10.1145/3290605.3300514 Accessed 15 feb 2024
Tauscher J-P, Schottky FW, Grogorick S, Bittner PM, Mustafa M, Magnor M (2019) Immersive eeg: evaluating electroencephalography in virtual reality. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 1794–1800. IEEE
Valve Corporation: SteamVR. https://store.steampowered.com/app/250820/SteamVR/?l=english. [Accessed 04-03-2025] (2014)
Villar A, León C (2023) An Evaluation on the Effectiveness of Virtual Tutorials for Training Emergency Professionals

Villar A, León C (2023) User Performance Analysis in Guided and Non-guided Stressful Virtual Reality Scenarios. https://doi.org/10.54941/ahfe1004038

Vosinakis, S.: The use of digital characters in interactive applications for cultural heritage. In: Applying Innovative Technologies in Heritage Science, pp. 109–137. IGI Global, USA (2020)

Wang J, Shi R, Zheng W, Xie W, Kao D, Liang H-N (2023) Effect of frame rate on user experience, performance, and simulator sickness in virtual reality. IEEE Trans Visual Comput Graphics 29(5):2478–2488

Yousef AMF, Chatti MA, Schroeder U (2014) The state of video-based learning: a review and future perspectives