UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Psicologia Generale
Brain, Mind, and Computer Science Ph.D. Course
Neuroscience, Technology, and Society curriculum
XXXV cycle

Coping with the inheritance of COVID-19: the role of new interactive technologies to enhance user experience in different contexts of use

Supervisor: Luciano Gamberini
Co-supervisor: Mauro Conti

Ph.D. Student: Alice Bettelli

Academic Year: 2022/2023

To Dad.

Table of Contents

Abstract **9**

1. Introduction **11**

1.1 The project: studies and contributions 13

1.2 Key concepts of Virtual Reality (VR), Augmented Reality (AR), and Interactive Streaming Services 16

1.3 HCI, Human-centric design, and the importance of mixed methods 18

2. Study 1 - Training in immersive VR **20**

2.1 Introduction 20

2.2 CO-design of the immersive VR experience 22

2.2.1 Background 22

2.2.2 Materials and Methods 24

2.2.2.3 Gathering the requirements: affinity diagram sessions 26

2.2.2.4 Planning the virtual experience: brainstorming session 27

2.2.3 Results 28

2.2.3.1 First affinity diagram session: risk identification 28

2.2.3.2 Second affinity diagram session: facing the emergency 29

2.2.3.3 Brainstorming session: finalization of the virtual simulation design 31

2.2.4 Implementation of the virtual simulation 33

2.2.5 Discussion 34

2.2.6 Guidelines for virtual simulation design of flood emergency 36

2.2.7 Conclusions 38

2.3 Evaluation of immersive VR experience 39

2.3.1 Background 39

2.3.2 Preliminary experts' evaluation and implementation of suggestions 40

2.3.3 Description of immersive VR experience..41
2.3.4 Evaluation of immersive VR experience with target users..46
2.3.4.1 Materials and method 46
2.3.4.2 Participants...51
2.3.5 Analysis and Results 52
2.3.5.1 Knowledge of flood emergencies...52
2.3.5.2 Perceived preparedness and attitudes towards flood emergencies 55
2.3.5.3 Commands of immersive VR experience: subjective and objective learnability....55
2.3.5.4 Performance in the immersive VR experience..56
2.3.5.5 Users' opinions about the immersive VR experience 57
2.3.5.6 User preferences about training in flood emergencies..59
2.3.6 Discussion..60
2.3.7 Conclusion...62

3. Study 2 - Shared watching experience and interactive streaming service.......................63
3.1 Introduction 63
3.2 Background...64
3.2.1 Interactive digital narratives...64
3.2.2 Computer-mediated group decision making 66
3.3 The Study...68
3.3.1 Experimental design 68
3.3.2 Task 68
3.3.3 Episode structure 69
3.3.4 Measure 70
3.3.4.1 Objective data...70
3.3.4.2 Self-report data...70
3.3.5 Experimental procedure..72
3.3.6 Participants..73

3.4 Analysis and results..73
3.4.1 Distribution of the narrative choice..73
3.4.2 Preference for the ideal path..74
3.4.3 Distribution of comments produced during the episode..........................75
3.4.4 Comments production between narrative crossroads and scene phases..........76
3.4.5 Contents of comments at the narrative crossroad.....................................76
3.4.6 User experience evaluation questionnaire...78
3.4.7 Group psychosocial effects...79
3.5 Discussion..81
3.6 Guidelines for the design of interactive streaming services.........................82
3.7 Conclusion...83

4. Study 3 - Retail and 3D web and AR..84
4.1 Introduction...84
4.2 Rethinking the shopping experience: explorative study on 3D web and AR for displaying and selling goods...86
4.2.1 Background..86
4.2.2 Materials and methods...87
4.2.2.1 Definition of user categories and requirements.......................................87
4.2.2.2 Participants...87
4.2.2.3 Semi-structured interviews..88
4.2.2.4 Procedure...89
4.2.3 Results..89
4.2.3.1 Retailers' experience with e-retail..89
4.2.3.2 Trade fairs operators' experiences..91
4.2.4 Discussion...92
4.2.5 Design Guidelines..93
4.2.6 Conclusion...94

4.3 Customer experience with 3D virtual objects in AR shopping 95
4.3.1 Background 95
4.3.1.1 The potential of AR in shopping experience 95
4.3.1.2 Interaction with Virtual 3D Objects 95
4.3.2 Data collection 97
4.3.2.1 Experimental Design 97
4.3.2.2 Equipment and technical set up 98
4.3.2.3 Purchasing scenarios 99
4.3.2.4 Measurements and materials 100
4.3.2.5 Procedure 104
4.3.2.6 Participants 105
4.3.3 Analysis and Results 107
4.3.3.1 self-report data 107
4.3.2.1 Behavioral data 112
4.3.4. Discussion 115
4.3.5. Guidelines for AR purchase applications of household products. 117
4.3.6 Conclusion 117

5. General conclusions 119

Reference 123

Abstract

The COVID-19 pandemic has upset the habits of people and various sectors of society, including training, entertainment, and retail. These sectors have been forced to adapt to abnormal situations such as social distancing, remote work, and online entertainment.

The pandemic has significantly transformed the training field, leading to the closure of many in-person instruction centers and a shift toward online education courses, which can be less effective. In addition, the entertainment industry has been heavily transformed by social distancing, resulting in the cancellation of many live events and the closure of several cinemas. This has increased demand for online entertainment options, such as streaming services and virtual events. Finally, the restrictions imposed by the COVID-19 pandemic substantially impacted physical stores and fairs, suspending exhibitions for more than two years. This has further driven consumers to rely on e-commerce to fulfill their purchasing and companies to increasingly take advantage of new technologies such as augmented reality. In this suddenly disrupted scenario, new technologies have the potential to fill the gap generated by the pandemic, functioning as an interactive bridge to connect people.

This Ph.D. thesis explored the potential of interactive technologies in mitigating the challenges posed by the COVID-19 pandemic in various contexts of use in the above-mentioned areas. Specifically, three studies were investigated by conducting different data collection using a mixed approach in the Human-Computer Interaction field. The first study focused on the research of immersive virtual reality training, with a particular interest in flood emergencies, a growing phenomenon. The goal was to implement engaging and efficient training for citizens that live near rivers through a human-centric design approach. The second study explored innovative ways to improve social interaction and collaboration in the entertainment sector, highlighting guidelines for the design of shared streaming experiences. In particular, three different communication modalities were investigated during group viewing of an interactive film on a streaming platform. Finally, the third study focused on the retail sector. On the one hand, the focus consisted of understanding which aspects of the 3D web and AR technology are helpful for supporting small businesses and trade fairs. On the other hand, the focus was to investigate how to support consumers during an AR shopping experience when interacting with 3D virtual products of different sizes.

Overall, this project provides suggestions and guidelines for designing systems that can both increasingly connect people at a distance and offer new hybrid worlds. In addition, this project expands state-of-the-art related to interactive technologies and offers generalizable results outside

the crisis created by COVID-19. These technologies, now increasingly integrated into everyday life, can be a tool for empowerment and resilience, improving people's lives.

1. Introduction

The COVID-19 epidemic has affected people worldwide, especially regarding their mental and physical health. In some countries, such as Italy, the government ordered a lockdown to restrict the pandemic spread and safeguard vulnerable populations (e.g., elderly and immunodeficient people) (D.P.C.M., 2020). Due to preventive constraints and fear of infection, people started to avoid large crowds. In this framework, urban systems have shown their vulnerabilities and fragilities, not having always been able to reorganize public spaces promptly to cope with social distancing (Mahima et al., 2022). Many companies, for example, have had to shut down, which has had an evident impact on the country's economy (Maital et Barzani, 2020) and the impoverishment of people's everyday experiences in terms of entertainment and recreation, and in terms of training and education (Kumar et al., 2021). In this limitation of individual freedom in non-work activities, interactive technology represented a way to improve our daily activities and enjoy experiences more safely and inclusively.

In the pandemic context many training courses have been suspended or canceled and replaced by online training (Gemmanno et al., 2022) since it has become difficult to guarantee face-to-face training. To fill this gap, Virtual Reality (VR) has been widely employed to safely and realistically recreate risk environments and emergency training. Indeed, this relatively new technology presents some important advantages. On the one hand, the fact that people tend to respond to situations presented in VR as if they were real (Rovira, 2009) has made VR the perfect tool for studying how people naturally behave in emergencies. On the other hand, the fact that virtual simulations highly engage users and allow them to experience firsthand the consequences of their actions have meant that this technology could be effectively used to teach the correct procedures in risky contexts (Çakiroğlu & Gökoğlu, 2019; Chittaro & Buttussi, 2015). Among the possible emergency situations which could be explored in VR, river floods represent one of the most frequent and threatening climate events, and because of climate changes, they are expected to become even more frequent and intense (W.H.O.). Very few works have used VR to create virtual training to face them (Zaalberg & Midden, 2013; Sermet & Demir, 2019). More in-depth research on which contents should be included and how to present them properly is still lacking.

The entertainment sector is a further field strongly impacted by the pandemic emergency because of the social distances and the government's preventive measures. In this scenario, the number of subscribers to streaming platforms has grown (Parnami & Jain, 2021), and a substantial increase in the use of video streaming services among adolescents is pointed out (Fernandes et al., 2020). Nowadays, it is possible to have a group viewing online experience, thanks to specific extensions implemented in some streaming services (i.e., Teleparty). Computer-mediated communication while watching online episodes has been studied in the literature (Geerts, 2006; Weisz et al., 2007) but has not yet been applied to the field of interactive movies. This kind of entertainment could increase sociality among users and build active interaction, reducing in somehow the fatigue of the stay-at-home measures and overcoming social distance.

Finally, the pandemic restrictions have also affected physical stores and fairs, with all the events and exhibitions suspended for more than two years; this pushed even further consumers to rely on e-commerce (Jílková & Králová, 2021). Despite being very popular and convenient, the online shopping experience cannot fully resemble the experience one can have in brick-and-mortar shops. Indeed, the customer cannot manipulate and touch the product, thereby failing to appreciate its tiny details. However, novel technologies such as VR and Augmented Reality (AR) have the potential to overcome these limitations by compensating for lack of sensory stimulation (Hilken et al., 2022). Therefore, it is important to develop e-commerce platforms that leverage 3D web and AR technologies to enhance the user experience, not only for consumers but also for sellers. To this end, it is crucial to understand end users' needs. In addition, the individual experience of AR with new devices (e.g., HoloLens) could help avoid overcrowding and maintain social distancing (Billewar et al., 2021). Several studies on AR and retail have focused on business and economic aspects (Billewar et al., 2021), but little attention has been given to the aspects of Human-Computer Interaction (HCI) related to manipulation of 3D products.

In this scenario, rapid advancement in interactive technology should play a crucial role in response to the COVID-19 pandemic. The main purpose is to break down the barriers of the physical environment in various contexts of use by creating hybrid spaces that are more resilient and adaptable. This leads to more and more orienting towards an extended reality in which people actively participate in their experience.

1.1 The project: studies and contributions

The main purpose of this project was to understand how new technologies can help to reduce the weaknesses COVID-19 has brought out in different user scenarios, namely entertainment, training, and retail. In particular, the goal was to test different immersive and interactive technologies in order to increase sociability and cooperation in the entertainment field, keep the retail sector alive through the new hybrid purchase experiences, and maintain and enable emergency training even remotely.

Three main objectives were considered:

- Evaluating which characteristics of immersive VR training effectively teach people how to cope with flood emergencies.
- Identifying which communication system is the most suitable to improve cooperation and user experience during an online group vision of interactive TV series.
- Understanding which aspects of the 3D web and AR technology are considered helpful for supporting small businesses, trade fairs, and customer purchases.

This project represented an advancement in the areas of Human-Computer Interaction in these different contexts of use. In particular, the results provided insights on the efficiency of new immersive technologies as tools to enhance users' daily experience (i.e., shared watching experience in interactive streaming services, training in immersive VR, 3D web, and AR for product display and sales). Finally, the guidelines that emerged from studies are useful to the scientific community, companies, and organizations for the design of new interactive technologies in different contexts of use. In addition, this could lead to an advance for society, safeguarding people during periods of crisis when public health could be compromised.

Study 1: training in immersive VR

The first study aimed to implement an immersive virtual simulation for training and education in flood emergencies. In particular, simulation requirements are investigated using a participatory design approach involving different stakeholders to identify the design elements (e.g., contextual, fundamental, and interactive aspects) and learning contents and modalities (e.g., learning objectives and learning activities tasks). The process of interaction design to develop the virtual simulation passed through the following steps:

- co-design activities (i.e., affinity diagram sessions and brainstorming) involving stakeholders with expertise in various sectors (i.e., HCI, psychology, VE design, educational training) to provide the requirements of the aspects characterizing the virtual reality simulation.
- implementation of the virtual simulation resulting from the early stages of co-design.

- activities of experts’ evaluation of the developed scenario.
- re-design and implementation of the virtual simulation to integrate the experts' suggestions.
- final users’ evaluation of the virtual simulation.

Study 2: shared watching experience and interactive streaming services

The second study aimed to investigate a streaming group viewing experience by offering different communication systems to interact with other people online during the view of interactive media. In particular, the main research goal was to identify which communication system might be implemented during an online group relating to watching an interactive episode of a TV series.
In this regard, a between-subject study was carried out, where several user groups tested an interactive episode of a tv-series with different communication systems. The independent variable was the communication system, presented in 4 different levels: absence of a communication system (no support), a text chat, an audio chat, and a video-audio chat. Data collected included group dynamics, narrative choices made in the interactive episode (i.e., video recordings), and self-reported quantitative ratings about user opinions (i.e., questionnaires). The data regarding efficiency, goal achievement, user experience evaluation, and communication quality were compared.

Study 3: Retail and 3D web and AR

The third study consisted of understanding which aspects of the 3D web and AR technology are considered helpful for supporting small businesses and trade fairs. To accomplish this objective, a first data collection was carried out involving two groups of stakeholders, namely: retailers and trade fair operators. Specifically, through a series of semi-structured interviews the factors that motivate small businesses to launch and manage online shops, along with the obstacles that they find in this activity, were explored. Moreover, it was investigated whether and how they perceive that 3D web could be helpful for them. Finally, the motivation of trade fair operators to adopt AR technology to improve product display and to maintain continuity during lockdown periods was investigated. A second data collection was conducted to understand how to support user interaction with 3D products during an AR shopping experience with HoloLens devices. In particular, the main research goal was to investigate the users’ preferences related to the manipulation of different sizes of 3D virtual products during a shopping experience in AR. The research involved the target user (i.e., purchaser) and adopted a within-subjects design. The independent variables were the type of manipulation (automatic handling vs. free handling) and the size of the products (large objects vs. small objects); (2x2). The data collected relates to user behaviors (e.g., number and time of manipulation) and the self-report data about the user experience during the purchase.

Next, we will introduce some key concepts to understand better the interactive technologies we will discuss in this project. Instead, below is information related to the research field of interest and the methodological approach used. This will then be followed by chapters relating to the lines of research previously explained. Finally, general discussions will be presented.

1.2 Key concepts of Virtual Reality (VR), Augmented Reality (AR), and Interactive Streaming Services

To understand interactive technologies such as VR and AR is essential to refer to three key concepts: immersion, sense of presence, and interaction (Andersen & Thorpe, 2009; Slater, 2009; Sundar, Xu, & Bellur, 2010). Specifically, immersion relates to the amount of senses stimulated, interactions, and reality's similarity with the stimuli used to simulate environments. This feature can vary according to the properties of the technological system used to convey the simulated environment (Slater, 2009; Freina & Ott, 2015).

The concept of immersion is particularly useful in classifying VR systems into three macro categories (Bamodu & Ye, 2013): i) non-immersive/low-immersive systems that allow users to view 3D virtual environments through a 2D monitor; ii) semi-immersive systems that typically include a high-performance graphics processing system coupled with a large monitor or projector; and iii) immersive systems that completely isolate the user from the real world through a Head Mounted Display (HMD).

Presence, the second key concept, refers to the extent of which users feel actually in the virtual place (place illusion), along with the reliably and effectively coming from it to perform certain actions (plausibility) (Slater, 2009). In VR and AR environments, the sense of presence is important in acting behaviors like the real ones (Botella et al., 2005; Juan et al., 2005; Bretón-López et al., 2010; Wrzesien et al., 2013).

Finally, the third key concept is interactivity, which can be defined as the extent to which media allow the user to exert influence over content and/or form (Steuer, 1992).

This last key concept is useful in understanding interactive streaming services. We can define them as an evolution of social television: the user can not only actively take part in the story by making narrative decisions that influence the plot, but can do so from the comfort of home, synchronizing with other users the media he/she is watching on a streaming platform (e.g., Netflix) and discussing the show via real-time text chat as if they were all sitting together (Mulla, 2022).

Defining VR (and, secondarily, AR), on the other hand, is more complex, as many authors have made heterogeneous proposals, with the result that unified and precise definitions are still lacking (Kardong-Edgren et al., 2019).

In this dissertation, we will refer to the following definitions of VR and AR: "*VR...isolates the perceptive channels of the subject by "immersing" him sensorially in the three-dimensional world generated by the computer. Immersion is made possible by a display and sound diffusion device which isolates the user from the outside world...* " (Pallavicini, 2020; p.32).

"*AR can be defined as a …technological system in which virtual objects are added to the real world in real-time during the user's experience*" (Cipresso et al., 2018).

To better understand these technologies, it may be useful to frame them within the so-called 'Reality-Virtuality Continuum' proposed by Milgram and Kishino in 1994 (Figure 1.1). At one end of the continuum, we find reality (physical word). We then find the so-called 'mixed reality' which includes augmented reality and augmented virtuality. In the former, there are virtual objects that are superimposed on the physical world. However, the real element is still predominant. In augmented virtuality, on the other hand, real elements are placed within a virtual environment. At the opposite extreme, we find virtual reality which defines environments consisting solely of virtual objects.

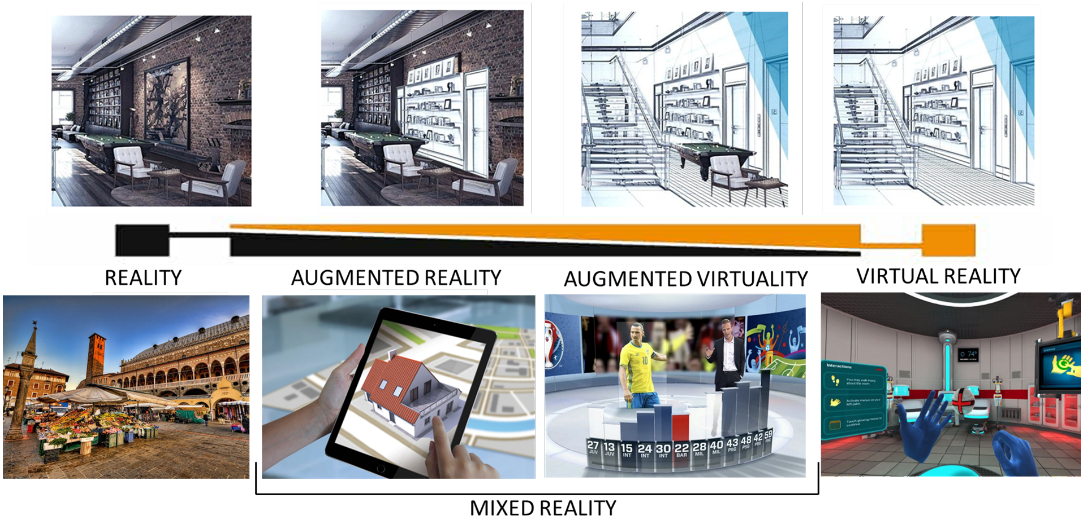

Figure 1.1: The Reality-Virtuality Continuum.

1.3 HCI, Human-centric design, and the importance of mixed methods

Human-Computer Interaction (HCI) is a multidisciplinary research field that draws on computer science, psychology, cognitive science, and organizational and social sciences to understand how people use and experience interactive technology. Technology is ubiquitous in modern life, and this multidisciplinary field is continuously evolving and shaping by new research and technological advancements. One problem with technologies could be that sometimes they can be extremely complex to use, forcing users to spend high effort in learning how to exploit them. For this reason, almost 25 years ago, the technology design started shifting from a top-down approach to a user-centered vision, ergo including users in the process (Preece et al., 2004). Human-centric design is an approach to product design that focuses on target users' needs, behaviors, and experiences. This approach involves target users in the design process and aims to create products that are intuitive, easy to use, and meet the needs of the user. It is rooted in principles such as usability and user experience. In literature, the most authoritative definition of usability is provided by the International Organization for Standardization (ISO, 2018): "The effectiveness, efficiency, and satisfaction with which specified users achieve specified goals in particular environments" (Boehm, Brown, and Lipow, 1976). The concept of "user experience" (UX) refers to how a product "behaves" and is used by people, encompassing all aspects of the interaction user-product (Nielsen & Norman, 2014). n the evaluation of usability, the primary focus is on the performance and fluid interaction, while UX takes into consideration desirable (e.g., satisfaction, engagement, enjoyment) and undesirable (e.g., unpleasantness, frustration, boredom) aspects of the interaction (Sharp Preece, & Rogers, 2019). The term "User Experience" was coined to describe what a person feels when interacting or imagining interacting with a prototype, a final product, a system, or a service. The International Organization of Standardization defined UX as: "a person's perceptions and responses that result from the use or anticipated use of a product, system or service" ISO (2010).

HCI has often used a mixed approach, which refers to using multiple research methods and techniques to gather data and insights about a particular user interface or interaction design. This approach combines the strengths of both quantitative and qualitative methods to provide a more comprehensive understanding of users' needs, behaviors, and attitudes toward a system. For example, quantitative questionnaires are used to explore users' backgrounds and gather information about their expertise and familiarity with technology (Wohlin et al., 2006; Issa & Isaias, 2015). They can also be administered retrospectively to investigate aspects of usability and user experience (UX) after the interaction with a product has taken place (Kortum & Bangor, 2013) or at the end of each task to inquire about specific aspects and functions (Wohlin et al., 2006). While questionnaires can collect a large amount of data quickly, they may not capture other relevant information that arises during the interaction, such as specific issues or problems not considered by the questionnaires (Van Elzakker et al., 2008). On the other hand, interviews are qualitative methods that can provide valuable insights into user feedback, motivations, or preferences. They are also commonly used for UX evaluation and can be conducted at any phase of the product life cycle, similar to questionnaires (Bevan et al., 2018). However, interviews require more time as they often involve audio and/or video recording for subsequent transcription, coding, and analysis (Toraldo et al., 2018), making them a more time-consuming method (Bartholomew et al., 2000).

Moreover, qualitative methods such as affinity diagramming (AD) and brainstorming (BS) are mainly used in the early stages of the interactive design process. AD allows groups of users to generate large amounts of qualitative data (e.g., ideas, opinions, issues) and to organize them. In particular, participants are asked to discuss the ideas generated and to group them in agreed categories organized hierarchically (Hartson & Pyla, 2012). BS is a technique utilized to support creative thinking and the generation of a wide range of ideas on a specific question or a peculiar problem. AD and BS are techniques variously recommended for data analysis, idea generation/planning, and consensus-based decision-making. Finally, other methods rely on objective data, such as observations and log data which can be collected during participants' interaction with the technologies and other users. Particularly, the video analysis method allows the observation of user behavior during interactions between individuals and technological artifacts within specific contexts. This method utilizes the support of video and audio recordings (Blackler et al., 2018) and can provide rich, detailed information regarding the usage context of a product, highlighting issues that may not be evident through other methods (Millen, 2000). In contrast, log data is collected through sensors and refers to the records of users' interactions with a computer system. These records can include timestamps, the actions the user performed, and any input or output data associated with those actions. For example, by identifying performance issues through the analysis of log data, researchers can work to improve the efficiency and effectiveness of the system.

In conclusion, mixed methods research has been shown to be effective in HCI as it can help researchers to overcome the limitations of using only one method. By triangulating data from multiple sources, mixed methods can provide a complete picture of the user experience and identify design information to support the interaction (Toraldo et al., 2018). In addition, the target user should be involved using different evaluation methods or adjusting evaluation criteria to account for the specific context of use. In fact, it is important for UX researchers and designers to consider the contextual factors that may influence the evaluation of UX and tailor their evaluation methods accordingly.

2. Study 1 - Training in immersive VR

2.1 Introduction

In recent years, Virtual Reality (VR) has become a popular tool for recreating situations that are difficult to access or investigate in the real world, such as emergency contexts and risky environments (Freina & Ott, 2015). VR simulations are so realistic that people tend to respond to them as if they were real (Rovira, 2009), making it an effective tool for studying how people behave in emergencies and the levels of stress they experience (Gamberini et al., 2003; Ronchi et al., 2015). Compared to traditional methods like manuals or video recordings, virtual trainings have been found to be more efficient, particularly when complex operations have to be taught (Chao et al., 2017).

Virtual reality can be classified based on the level of immersion provided by the technology, where immersive virtual reality (IVR) involves our senses more than other types of VR (Morganti & Riva, 2006). Researchers have shown that IVR is more engaging and provides a greater sense of presence than traditional displays (Chittaro & Buttussi, 2017). Moreover, VR simulations for emergency support aim to provide users with resources and knowledge to implement functional coping strategies in high-risk situations, where traditional approaches like slides have been found to be ineffective as they do not provide individual feedback or involve the user emotionally (Feng et al., 2018). Virtual training has also been considered a solution in contexts where traditional face-to-face approaches were not feasible due to COVID-19 confinement, provided there were technological means, digital skills, and attitudes to change (Sastre-Merino et al., 2020). Furthermore, the availability of affordable VR headsets like the Oculus Quest 2 has made IVR experiences accessible to the wider public (LaValle, 2019).

VR technology has been successfully used for educational purposes in the emergency field, such as teaching the correct procedures to implement in risky situations, usually through immersive serious games (Çakiroğlu & Gökoğlu, 2019; Chittaro & Buttussi, 2015). Various studies have explored the potential of VR for simulating different types of emergencies, including earthquakes (Tarnanas & Manos, 2001; Li et al., 2017), wildfires (Benvegnù et al., 2019; Kinateder et al., 2014) and military or terrorist attacks (Shendarkar et al., 2008). Recently, among the different types of emergencies that can be supported by immersive VR experiences, climate change such as typhoons (Ke et al., 2019) and floods (D'Amico et al., 2022, Mol et al., 2022; Fujimi and Fujimura, 2020) are gaining in importance due to climate disasters. Floods are a hydrogeological instability phenomenon that frequently happens in Italian territory[1,2]. Floods related works already exist and consider different aspects of this phenomena. For example, Fujimi and Fujimura (2020) proposed a VR application focused only on the promotion of evacuation, while Mol and colleagues (2022) used virtual reality to increase risk perception, coping evaluation, negative emotions, and harm reduction behavior through a simulated flood experience. Differently, D'Amico and colleagues developed a non-immersive serious game for flood emergency training showing a significant increase in self-efficacy and safety knowledge after the VR experience (D'Amico et al., 2022). Therefore, preliminary studies seem to support the potential of using VR for flood safety.

However, previous research has not utilized co-design methodologies to develop a VR simulation that can enhance the quality of life of individuals living near a river by providing them with concrete information to better manage flood emergencies. To address this gap, this study outlines the virtual educational experience "Safer Water", focusing on the outcomes of a series of co-design sessions that involved different stakeholders. Particularly, experts from various fields worked together to create immersive and interactive virtual scenarios and, once the virtual reality simulation was implemented, target users were involved to evaluate the experience and verify its effectiveness.

The next sections will describe the interaction design process related to the "Safer Water" experience as follow: in the first part (sections 2.2 Co-design of the immersive VR experience[3]) the co-design activities carried out with the experts related to the discovery of requirements for interaction, scenario design, and prototype implementation will be described; in a second part (sections 2.3 Evaluation of immersive VR experience) the preliminary evaluation of the simulation with experts, feedback and related redesign of the simulation, and evaluation of the experience with target users will be reported.

[1] https://www.legambiente.it
[2] https://www.isprambiente.gov.it
[3] The work presented in this section has been published in Gamberini, L., Bettelli, A., Benvegnù, G., Orso, V., Spagnolli, A., & Ferri, M. (2021). Designing "Safer Water." A Virtual Reality Tool for the Safety and the Psychological Well-Being of Citizens Exposed to the Risk of Natural Disasters. *Frontiers in psychology*, *12*, 674171.

2.2 CO-design of the immersive VR experience

2.2.1 Background

The purpose of previous works mentioned in the introduction (Kinateder et al., 2014; Benvegnù et al., 2019; Tarnanas and Manos, 2001; Hagita et al., 2020; Shendarkar et al., 2008; Ke et al., 2019; Fujimi and Fuji mura, 2020) is to observe how people react to emergencies or to teach how to cope with them, there is still limited research showing how VEs should be designed to convey appropriate social and psychological insights to participants. In particular, it is not easy to find a balance between the need to represent the emergency in a realistic way - allowing users to become emotionally activated - and the need to clearly suggest the individual and social behavior that people should have in the specific situation represented. In addition, there is a certain disparity in the study of the different risk contexts from a psychological point of view. For example, situations such as fire emergency have been extensively investigated in VR, focusing on evacuation procedures (Kinateder et al., 2014), and training (Çakiroğlu & Gökoğlu, 2019; Williams-Bell et al., 2015), while for other kind of emergencies, such as river floods, these domains are still relatively less considered. River floods are one of the most frequent and threatening climatic events and, due to climate change, are expected to become even more frequent and intense (OMS, n.d.). River floods have an increasingly devastating impact on communities and territories, reducing the safety and quality of life of people living near a river (Mason et al., 2010).

Few works on this topic have used VR to study flood-related behavior. Among these, a recent study (Fujimi & Fujimura, 2020) used immersive VR to test the effectiveness of interventions designed to encourage evacuation decisions from flash floods. Results showed that participants responded to the environmental and social inputs provided and that the effectiveness of flood evacuation interventions can be empirically examined using VR simulations (Fujimi & Fujimura, 2020). In the work of Zaalberg & Midden (2013), participants attended a virtual simulation of a levee breach on a desktop screen, showing (in a post-simulation evaluation) a higher motivation to evacuate, information seeking and a stated preference to purchase flood insurance compared to the other methods tested. Except for these researches and a few other exceptions (Sermet & Demir, 2019), most studies on this topic have focused on how to construct VEs that allow non-experts to visualize numerical simulations of changes in hydrological information (Lai et al., 2011; Leskens et al., 2017; Liu et al., 2018; Macchione et al., 2019). In some cases, these 3D environments have been employed for flood risk communication and used in public meetings, festivals and workshops to raise awareness of the risks associated with these extreme events (Lai et al., 2011; Skinner, 2020). However, most of these studies focus on the 'technical' construction of the scenario (e.g. the development of the numerical model, methods to improve the visual quality or the final rendering of the scenario), while more in-depth research on what content to include and how to present it properly is still lacking. Furthermore, as far as we know, none of the previous works have used co-design methodologies to create a VR simulation to improve the quality of life of citizens living near a river by providing them with concrete information to better face the flood emergency.

To fill this gap, in the present work, we describe the result of a series of co-design sessions aiming to bring experts from different fields to collaborate in setting up immersive and interactive virtual scenarios. Indeed, co-design has been already effectively applied in other areas related to VR experience. It was mainly used in mental health domain, to create virtual scenarios to facilitate psychological, cognitive and behavioral interventions for dementia, anxiety disorder, eating disorders, pain management (Tabbaa et al., 2020), and fear of public speaking (Flobak et al., 2019). Co-design techniques have also been effectively applied to develop VR simulations and serious games to promote physical activity (Boger et al., 2018; Eisapour et al., 2020) and to enrich user experience in cultural and naturalistic site (Bettelli et al., 2019). Although co-design does not seem to have been applied in the field of virtual simulation of emergencies, it was still used to design virtual training to manage stressful situations, such as training for police forces in the field of close protection (Lukosch et al., 2012), reentry training for incarcerated women (Teng et al., 2019) and alcohol resistance training for adolescents (Lyk et al., 2020). From a methodological point of view, these works are very heterogeneous both in terms of the co-design techniques used and the type of stakeholders who participated in the activities. Regarding the first point, several techniques were applied in the initial stages of requirements collection and/or experience planning, including affinity diagrams (Bettelli et al., 2019), brainstorming (Lukosch et al., 2012; Lyk et al., 2020; Tabbaa et al., 2020), focus group sessions (Boger et al., 2018; Eisapour et al., 2020) and interviews (Lukosch et al., 2012; Teng et al., 2019). Regarding the second point, these activities mainly involved final users (Lukosch et al., 2012; Teng et al., 2019) and/or experts in the specific application domain, for example kinesiologists for applications to promote physical activity (Eisapour et al., 2020) or managers of private security companies to design virtual training for police forces (Lukosch et al., 2012), while only few studies also involved experts in new technologies and HCI (Bettelli et al., 2019; Tabbaa et al., 2020). The co-design activity helped to obtain valuable feedback from experts (Eisapour et al., 2020), and to identify design recommendations for the specific application domain (Boger et al., 2018; Eisapour et al., 2020) even if it is not always easy to mediate between the different stakeholders's interests, risking slowing down the design process (Lukosch et al., 2012). The adoption of co-design methodologies has allowed the creation of detailed and customized virtual experiences. On this basis, we expect that a co-design approach can be successfully applied to the emergency context. Furthermore, as far as we know, none of the previous works have used co-design methodologies to create a VR simulation to improve the quality of life of citizens living near a river by providing them with concrete information to better face the flood emergency.

The general objective was to collect information useful for defining the VE contents and their translation into the virtual educational experience "Safer Water". Particular attention was paid to the spatial and temporal domain specification, focusing on events, effects, and possible behavioral responses in a situation where the embankment of a river was broken. The following sections will describe the methodologies used, the parties involved in the various activities, and the main results.

2.2.2 Materials and Methods

To create an immersive and interactive virtual experience that can increase the safety perception and risk awareness in people living in the proximity of a river, a co-design approach was adopted. By doing this, different experts (N= 11; Table 2.1) in several fields were involved in the design process and informed the requirements of the experience.

<table>
<tr><th>ID</th><th>Background</th><th>Stakeholders' professional expertise</th><th>Previous experience in simulation and VR</th></tr>
<tr><td>P1
P2
P3</td><td>Hydrology</td><td rowspan="2">1) Hydraulic and hydrological modeling;
2) Flood risk assessment and management;
3) Implementation of flood forecasting systems;
4) Management of forecasting and warning systems with the Civil Protection;
5) Citizen Observatories (especially P1; P2).</td><td rowspan="2">No familiarity with virtual reality, but experience in the field of simulations:
1) Implementation of flood simulation models;
2) Simulations of flood events and exercises for the evaluation of technologies and methodologies developed in the context of European Projects;
3) Organization and planning of various civil protection exercises.</td></tr>
<tr><td>P4
P5
P6</td><td>Hydraulics</td></tr>
<tr><td>P7
P8
P9</td><td>Psychology and Human-Computer Interaction</td><td>1) Interaction design, ergonomics;
2) Participatory design activities and users’ study;
3) Creation of storyboards for VR experiences;</td><td>Use of VR Simulations for scientific research purpose in different contexts (e.g., risk management and emergency situations, training and safety in the workplace, clinical area,</td></tr>
</table>

		4) Testing and evaluation of VR and other new technologies; 5)Persuasive technology.	naturalistic and cultural heritage, architectural, retail and product sales).
P10 P11	Design and implementation of VE	1) Creation of storyboards for VR experiences; 2) Architecture design of software in VR; 3) Creation of navigable 3D models optimized for VR and implementation of interactions.	

Table 2.1 Description of the stakeholders involved in the co-design activities: field, expertise and previous experience in simulation and VR.

They were experts in hydrology, hydraulics, psychology and Human-Computer Interaction, and design and implementation of VE (Table 2.2 describes expertise and previous experience with simulations and VR). These stakeholders took part specifically in the affinity diagram and brainstorming activities.

ID	Age	Gender	Education	Professional Experience (years)
P1	30	F	Master's degree in Engineer for the environment and the territory	17
P2	41	F	Ph.D. in Civil and environmental engineering sciences	5
P3	49	M	Ph.D. in Hydrodynamics and environmental modeling	11
P4	42	M	Ph.D. in Environmental hydronomy	6
P5	29	M	Master's degree in Civil engineering, hydraulic specialization	15
P6	35	M	Ph.D. in Hydraulic risk management	20
P7	54	M	Ph.D. in Experimental psychology	30
P8	29	F	Ph.D. student in Neuroscience, technology, and society; master's degree in clinical psychology	4
P9	30	F	Ph.D. student in Neuroscience, technology, and society; master's degree in neuroscience and neuropsychological rehabilitation	5
P10	32	F	Master's degree in Architecture	4
P11	36	M	Not graduated	8

Table 2.2 Description of the stakeholders' background: age, gender, instruction and professional experience.

2.2.2.3 Gathering the requirements: affinity diagram sessions

Two affinity diagram sessions were organized, considering two different moments relating to the spatial and temporal domain of the emergency situation. In particular, the goal of the first session was to collect information on the situation prior to the breakdown of a river embankment and on the possible danger indicators that signal with high probability an imminent flood. In comparison, the goal of the second session was to collect information on the emergency created by the rupture of the riverbank and on the consequent behaviors adopted by people. The affinity diagram methodology is widely used to generate, make sense and organize large amounts of unstructured, far-reaching, and apparently dissimilar qualitative data (Hartson & Pyla, 2012; Lucero, 2015).

First Session: Before the embankment breakdown

The first session focuses on the time frame before the breakdown of the embankment and involved eight participants and one conductor. The sample included experts in hydrology (P1, P2, P3), hydraulics (P4, P5, P6), and psychologists with expertise in Human-computer interaction and new technologies (P8, P9). This heterogeneity made it possible to grasp every perspective related to the development of the scenario. The activity took place in a setting that favored the production and elicitation of ideas. The participants gathered in a room and arranged in a semicircle around the conductor. The conductor was a psychologist with prior experience in conducting affinity diagrams (P7). In a first phase, the conductor created a convivial atmosphere, introducing the activity, and the participants presented themselves and their expertise. Then, each participant had the task of answering the focus questions posed by the conductor relating to the situation prior to the collapse of the embankment. The questions were intentionally structured in a generic way, in order not to influence or limit the participants in producing the contents. More specifically they were: "What aspects of a flood emergency situation should be considered in a virtual simulation?" "How could these aspects be represented?". The conductor favored the emergence of spontaneous ideas about the topic analyzed, encouraging the use of creative and non-logical thinking. All participants produced a series of ideas and wrote every single idea on a different card (previously provided by the conductor) in the most concise and clear way possible. The ideas that emerged were subsequently read aloud by the conductor and discussed one by one among all the participants. The purpose of discussing ideas was to assign each of them to broader groups based on criteria similarity in order to organize what emerged into unanimously agreed groups (Figure 2.1). To do this, the generated cards were placed on a white wall, to help the participants in the process and to have an overall view. During this phase, the conductor encouraged people to contribute their points of view. The groups obtained were then unanimously labeled: the participants read the contents of each group and wrote the name that best represented each category on a new card. Furthermore, some categories were grouped into named macro-categories. Finally, the categories were hierarchized and related to each other. During the activity, the participants were involved and proactive, producing various ideas and actively participating in the discussion. The session lasted about 3 hours.

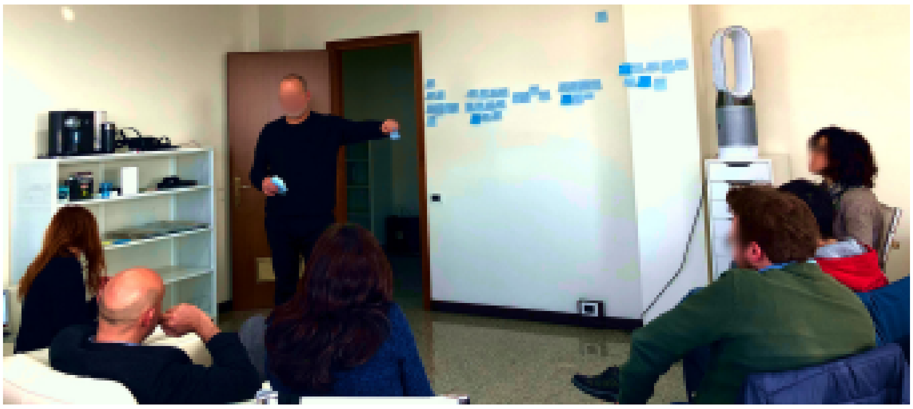

Figure 2.1: The first affinity diagram session.

Second session: After the embankment breakdown

The sample of the second session consisted of seven participants (and one conductor), including expert in hydrology (P2), experts in hydraulics (P4, P5, P6), psychologists with expertise in Human-computer interaction (P8, P9), and a designer of virtual environments (P10). The same conductor of the first activity, a psychologist with prior experience in conducting affinity diagrams (P7), supervised the session. The activity concerned the temporal window following the break of the embankment. The setting and the procedure employed were the same of the affinity diagram carried out in the first session. The same formulation of the key questions was adopted but focused on the post-break situation. The ideas that spontaneously emerged in response to the conductor's focus question were discussed, organized, and hierarchized, similarly to what was done previously. The conductor managed the entire session by moderating the discussion among the participants for a total duration of 3 hours.

2.2.2.4 Planning the virtual experience: brainstorming session

The results obtained from the activities were discussed and analyzed for the realization of the virtual reality scenario proposals. To define the experience, a brainstorming session was done with one expert in Human-computer interaction (P8) and the two VE designers (P10, P11). Another expert in Human-Computer Interaction (P9) moderated the session. This co-design technique allows to involve participants in the generation of solutions in an informal and relatively unstructured way, with a judgment-free discussion (Dix et al., 2004). The available data from the two affinity diagrams were analyzed and translated into operational scenarios, considering the diversity of requirements that emerged. In particular, during a first phase most of the aspects highlighted during the previous activities were selected, giving particular importance to the contents that emerged as priorities in the hierarchy phase, and the ideas that emerged in both affinity diagrams. Then, the information was organized into a coherent narrative. Regarding the interaction with virtual space, different techniques and methods were considered by the same participants, taking into consideration previous studies, until a unanimous agreement was reached. The brainstorming lasted 4 hours.

2.2.3 Results

This section will illustrate the main results of the collaborative design sessions previously described.

2.2.3.1 First affinity diagram session: risk identification

Overall, in the first affinity diagram session related to the time before the embankment breakdown, several ideas were produced. The categorization and hierarchization of these ideas highlighted macro-categories, categories, and specific subsets (Figure 2.2). Two main macro-categories emerged, namely *Contents* and *Shape*, that users regarded as closely linked to each other for the experience.

The *Contents* macro-category was composed by the categories *Narration* and *Events*. In particular, *Narration* referred to the various narrative expedients that could be implemented to introduce the user to possible events (e.g., the possibility of bringing the user closer to the river in a scenario that includes a walk along the embankment; to be firefighter/civil protection that inspects the site along the embankment to check the state of alert). With *Events*, participants identified both the *Initial Scenarios* subset, that is the situations that may be represented in the virtual environment (e.g., the development of a river context close to flooding, the mouth of a river, and a river in flood with very high-water level), and the subset *Physical indicators of danger*. The latter referred to natural signals that indicate an impending flood or possible subsidence of the embankment (e.g., the presence of flooded manhole covers, whirling eddies, floating material). These were the indicators that all the participants considered characterizing elements of the scenario for the educational objective.

The *Shape* macro-category referred instead to the formal aspects of the simulation. It was characterized by two categories, respectively *Technical aspects* and *Artificial indicators*. The first one included the presentation modalities (e.g., the use of audio or video support within the scenario, multimodal feedbacks) and the possible methods that could be adopted to make the physical indicators of danger more evident (e.g., positioning a reference object in the river, such as a bridge, to show the height of the water level). Finally, in the *Artificial Indicators* category were collected the unnatural elements of alert (e.g., the presence of sirens, radio messages of civil protection, specific signage). Participants considered the two categories linked, because both provide elements to help the user identify the risk situation.

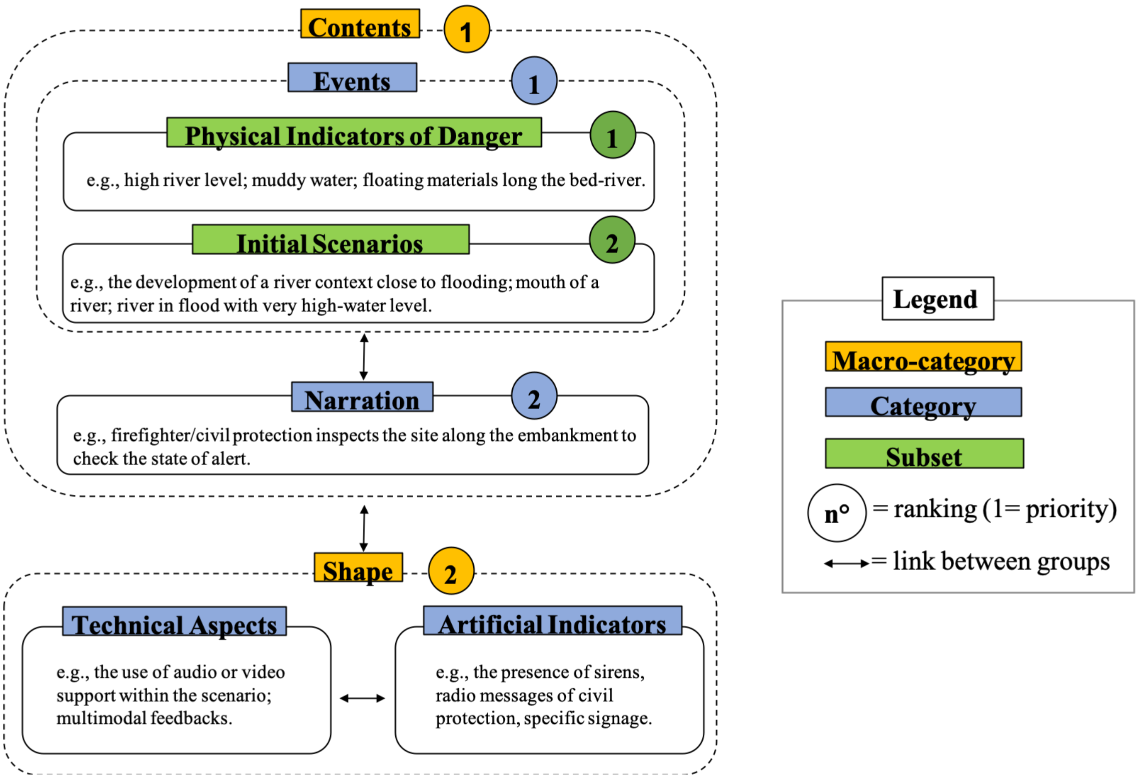

Figure 2.2: Schematic representation of the results of the first affinity diagram session.

2.2.3.2 Second affinity diagram session: facing the emergency

From the second session, focused on the emergency following the breakdown of the embankment, three distinct macro-categories were collected, namely *Actions*, *Architecture of the experience* and *Emotional outline*, with their specific categories (Figure 2.3). These groups were connected to each other by the users because they were closely related to the realization of the scenario.

The *Actions* macro-category includes the content aspects of the experience and focuses on the possible behaviors that could be taken during the emergency. These in turn have been categorized based on their correctness, with particular reference to the category *Things to do*, or rather the correct procedures to be adopted in the event of a break in the embankment, and *Things not to do*, or behaviors to avoid because they are considered dangerous or counterproductive during the emergency. The *Things to do* category includes, for example, actions such as avoiding flooded areas, reaching elevated places, calling for help. While, in *Things not to do*, various behaviors to

avoid emerged, such as recovering a car from at-risk places, using the elevator, climbing the high-tension pylons. Besides, a third category named *Uncertain decision situations* referred to ambiguous situations that characterized an emergency (e.g., how the user behaves when faced with a request for help). This macro-category was identified by the participants as the most important of the three, with particular reference to "things to do" and "things not to do" categories.
The *Architecture of the Experience* macro-category included *Technical aspects* and *Scenarios*. In line with what emerged in the diagram relating to the pre-flood situation, *Technical aspects* refers to the presentation modalities of the contents (e.g., the presence of feedback after the actions, the positioning of "false leads" to make the task more complex). The *Scenarios* contains possible specific situations that could be represented (e.g., a river context, user immersed in water up to the knee in a non-river context).
Finally, macro-category named *Emotional outline* signaled the importance of creating situations with a high emotional impact and classifies the elements that could influence the user's emotional state with *Positive elements* (e.g., the presence of civil protection vehicles) or creating a greater alarm whit *Negative elements* (e.g., screams, sirens).

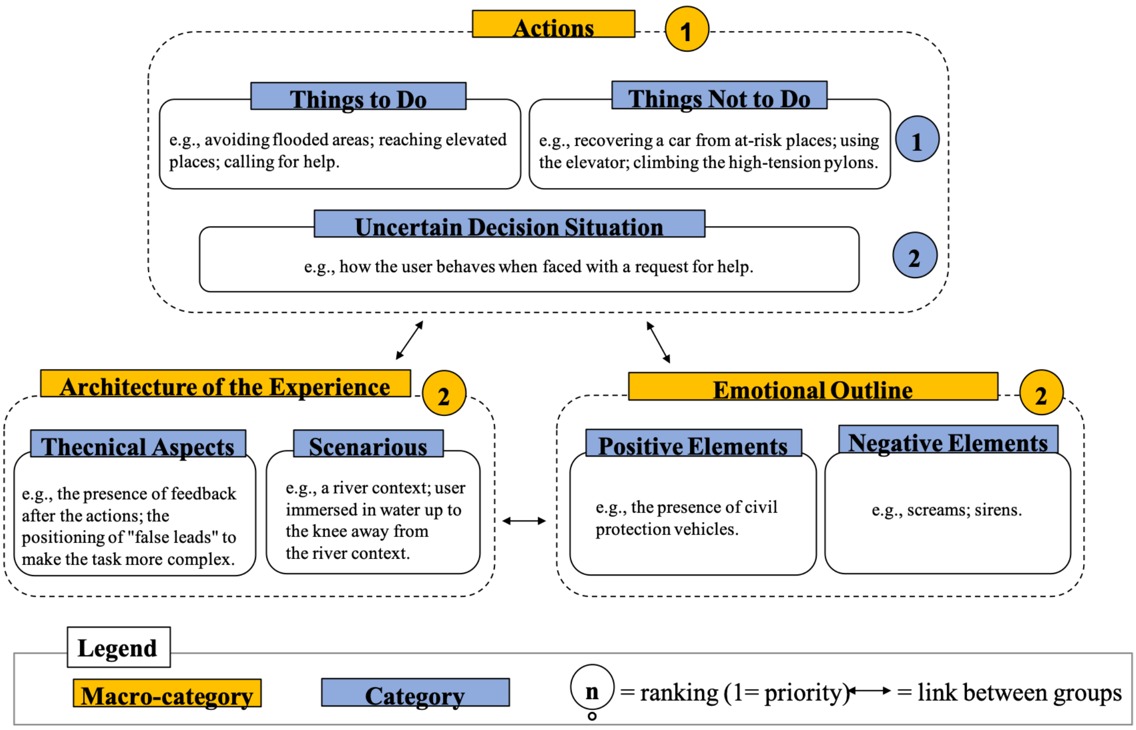

Figure 2.3: Schematic representation of the results of the second affinity diagram session.

2.2.3.3 Brainstorming session: finalization of the virtual simulation design

Based on what emerged from the affinity diagrams, the brainstorming session led to the identification of two educational and engaging scenarios: the first one focused on the time frame that precedes the breakdown of the embankment (i.e., preparedness) and the second one focused on the post-breakdown emergency (i.e., response). Both scenarios have been designed to be usable both continuously, with chronologically ordered events, and separately, by having the breaking of the embankment as the final salient element of the experience (first scenario) or as the initial one (second scenario). The focus of the experience was on the *Physical indicators of danger* before the embankment's breakdown and on potential *Actions* to be performed or avoided during the consequent evacuation, because these were the contents that emerged as priorities in the hierarchy phase.

The VE depicts a river area and the neighboring landscape (Figure 2.4A). To ensure the realization of an ecological scenario and a realistic simulation, the environment is characterized by the typical sounds of nature, the specific fauna that populates the river, and the presence of other citizens (e.g., fishermen).

In the first scenario, the focus is on the exploration of risk indicators in order to increase user awareness. For this purpose, five *Physical indicators of danger* of an impending flood or a possible bank failure have been selected and inserted in the VE: (1) flooded manholes covers (Figure 2.4B), (2) high river level and muddy water, presence of (3) floating material, (4) whirling eddies, and (5) "fontanazzi" (leakage of cloudy water near the embankment). Moreover, the scenario presents a bridge (*Technical aspects*) and typical *Artificial indicators* of the river context (i.e., hydrometers positioned along the bridge and along the staircase) in order to make the "slow" physical indicators of danger (i.e., the water level) more easily recognizable (Figure 2.4C). The *Narrative* invites the user to explore the river environment and identify potential risk signs of levee breakage. To do this, the simulation begins with the user (who impersonates a member of the civil protection service) in a parking lot near a river. A colleague informs him/her via a phone call (i.e., an audio directly into the headset) that (s)he is there to inspect the state of the embankment, by checking if there are any indicators of danger (the actual task of the first scenario). The user then follows the path, walking along the embankment between the two sides of the river, and at the end of the task another phone call invites him/her to join the colleagues in a specific point of the VE.

While the user approaches the indicated area, the narrative involves the subsidence of a point of the embankment (Figure 2.4D). The embankment breaks at a distance that does not physically overwhelm the user, but is sufficient to be able to observe the phenomenon, generating an immediate emergency. This causing the surrounding countryside to flood, with the water level rising (Figure2. 4E).

The second scenario is immediately subsequent the breakdown of the embankment and focuses on learning the correct behaviors to adopt during an emergency. In this case, the narrative requires the user that is in a countryside bank to evacuate the area and to stay safe. To do this, as the water level rises, voices in the distance yell to the user to save themselves. The scenario offers to the user various escape routes and provides different feedbacks depending on the choices made. Based on what emerged in the affinity diagram relating to the post-breakdown situation, in the simulation are implemented elements that lead to safety, saving himself/herself in a high and protected location (*Things to do)*, but also stimuli that could lead to dangerous conduct (e.g., climb a tree that is not resistant enough or take a car that is in at-risk place). The actions performed by the user will be followed by adequate immediate feedback according to the correct or incorrect behavior, from an educational point of view. Once a stable and safe position has been reached, the user must contact the emergency services (and not clog the lines by calling friends or family). Another aspect considered in the scenario is related to the *Emotional outline*, with the insertion of citizens running away (*Negative aspects*).

Finally, in both scenarios, elements of gamification (*Technical aspect*) were considered, such as the possibility of assigning scores to the user based on his/her actions or of running into a "game over" condition in the event of critical errors. These elements, typical of serious games, help in making the experience not only educational but also challenging. Also, they lighten the tone of the intrinsically dramatic situation represented and could be experienced as excessively stressful by users. The simulation ends when the user succeeds in calling emergency services.

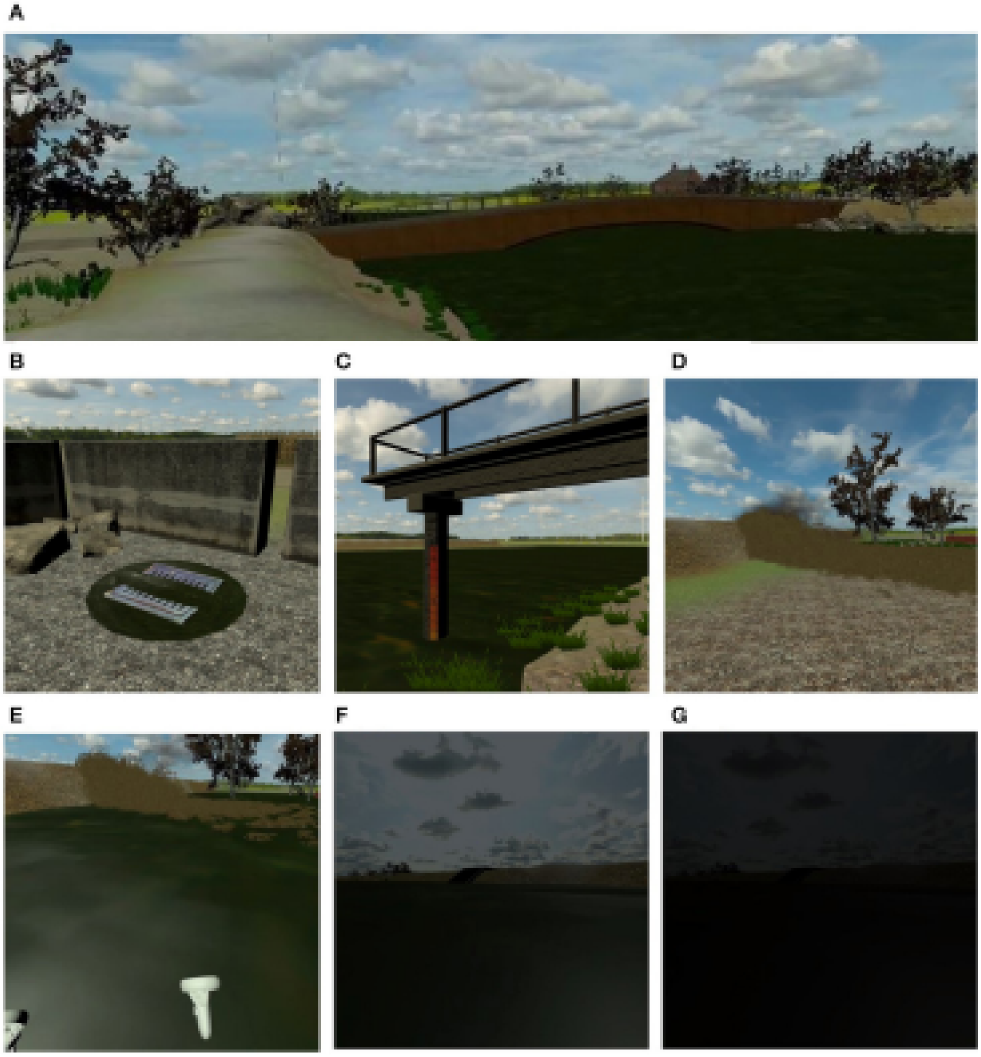

Figure 2.4. (A) A frame of the VE representing the initial part of the path along the river area. (B) A flooded manhole, one of the physical indicators of danger. (C) A hydrometer, an example of artificial indicator that facilitates the recognition of the high-water level. (D) The breakdown of the river embankment. (E) While the water level rises, (F) the user fails to reach a high point, and (G) the screen slowly fades to black.

2.2.4 Implementation of the virtual simulation

The two scenarios defined during the brainstorming activity were created using Unity (version 2020.1.17) by two experts of design and implementation of VE. In particular, the 3D models were built with blender 2.9. To make the simulation more realistic, the virtual environment was partially based on a river area in northern Italy known to be subject to frequent floods. The ambient audio used was recorded with a binaural microphone (Neumann KU100) in the field (i.e., river area). The resulting simulation has been configured to be usable through Oculus Quest 2 (resolution: 1832 x 1920px per eye; refresh rate: 90 Hz; Field of view: 100-degree est.; integrated speakers and microphone).

To select elements in the virtual environment (e.g., an indicator of danger) the “ray casting” method (using the trigger button of the Quest controllers) was employed. As locomotion modality, “teleportation” was selected (assigned to the thumb stick button). Both the chosen techniques are widely used in the field of VR and well received by users, in terms of ease of use, cognitive load and cybersickness (Loup & Loup-Escande, 2019; Nukarinen et al., 2018). Finally, on the basis of what emerged in the brainstorming session, an immediate feedback system was adopted in the simulation, similarly to (Chittaro & Buttussi, 2015). In case of adoption of an incorrect behavior or omission of a correct one, a negative feedback (with visual, vibrotactile and auditory elements) is provided, followed by a short recommendation. For example, if the user fails to reach a high point during the flooding, the sound of the water intensifies, the controller vibrates, the screen fades to black (Figure 4F and G) and then the user receives the message “Reach a location at a safe altitude”. Then, the user starts over from the point in the narrative where the mistake was made.

The simulation collects and logs the number of physical indicators of danger found, their type, when they were identified, the data relating to the actions performed, the path taken by the users (as a sequence of spatial coordinates), and the overall time spent in the scenario by the user.

2.2.5 Discussion

The purpose of the present study was to describe the co-design activities implemented to identify the contents of the VR simulation “Safer Water” for river flood emergencies. The focus was on designing virtual scenarios that would provide useful knowledge for users to deal with an embarkment breakdown with consequent flooding, with the ultimate aim of improving the quality of life of those who live near a river. To this end, experts with different theoretical backgrounds were involved in two affinity diagram sessions and a brainstorming session.

Overall, the co-design process we adopted for the creation of the VE has led to a series of advantages. Previous studies have shown how co-design techniques can be effectively used to design virtual scenarios related to specific situations, as it allows to obtain valuable information from experts in the reference sector (Eisapour et al., 2020). In our case, we have chosen to also involve HCI experts and VE designers from the beginning, a choice that was made only in a few of the previous works (Bettelli et al., 2019; Tabbaa et al., 2020). Involving this type of experts from the early design stages allowed us to focus not only on the contents of the VR environment (e.g., danger indicators of a flood, correct evacuation behaviors), but also on the technical aspects and interactive modalities (e.g., feedback, gamification elements), obtaining a global view of the various parts to be taken into consideration. We also found useful having carried out more than one affinity diagram session since it let us analyze the two different phases of the emergency separately. Indeed, emergency situations are complex phenomena, and having the possibility to break them into multiple phases with related activities was an aspect that facilitated the co-design process, giving the participants the time necessary to produce ideas and investigate in detail a series of elements related to a specific time frame of the natural disaster. However, it should be considered that if on the one hand doing multiple sessions allows a deeper analysis of the phenomenon, on the other it means risking not having the same participants in all the sessions.

Regarding the results of the co-design activities, two related virtual scenarios with different focuses emerged: a first scenario aimed at identifying the signs of a probable breakdown of the embankment (i.e., preparedness), and a second one concerned the behaviors to be adopted during the emergency (i.e., response).

Different narrative modes were associated with the different objectives. In the first scenario, an exploratory modality, based on researching and detecting the cues, was selected. In the scenario related to the situation following the embankment breakdown, an interactive-experiential modality was chosen: the user receives positive or negative feedback based on the correctness of his/her behavior in the VE.

Compared to previous literature on VR and river flood emergency, our results highlighted three fundamental aspects. First, the need to consider not only the emergency in progress but also the time frame before the disaster has clearly emerged. Instructing users on the environmental elements to pay attention to in order to estimate the probability of an embankment breakdown could be fundamental to increase their risk awareness and well-being. Indeed, such calamities are often perceived as sudden and unpredictable events, and we expect that providing citizens with the knowledge to grasp the signs that precede them can improve their perception of control over the environment and, consequently, their quality of life. However, previous works on the topic largely ignored this aspect by placing the user already in the emergency or shortly before its occurrence (Fujimi & Fujimura, 2020; Zaalberg & Midden, 2013).

Secondly, from the affinity diagram and the brainstorming sessions, a preference to present the indicators of danger in a natural and realistic way emerged in order to facilitate the transfer of the information learned into a real context. For example, the river's water level in our virtual scenario is already relatively high, an aspect that can be identified thanks to references such as bridges or hydrometers along the path. Previous works have instead preferred to artificially accelerate the time flow in the simulation, a system that allows researchers to show in a few minutes changes that occur over hours at the cost of diminishing the realism of the experience (Fujimi & Fujimura, 2020; Skinner, 2020).

Finally, the co-design activities made it possible to highlight which actions could be done and which should be avoided during the flood emergency, giving immediate feedback to the user. The possibility of letting users experience firsthand the consequences of their behavior in an immediate and vivid way was a design choice adopted in various studies on emergency and risk situations (Buttussi & Chittaro, 2017; Chittaro & Buttussi, 2015; Feng et al., 2018; Van Ginkel et al., 2020), but still little used in the works concerning specifically the flood emergency, more centered on allowing the users to visualize the disaster and not to actively act to face it. In particular, a preference for an immediate feedback system emerged from the co-design activities. The literature on the subject is not conclusive: in the context of VR emergency training, both immediate and delayed feedback have been argued to be effective to enhance trainees' preparedness (Feng et al., 2021). Some evidence in favor of deferred feedback was found with child users (Feng et al., 2021), while in the case of adults the question is still open. However, besides the timing, the way in which feedback is provided is also of considerable importance. In our case, it was designed not only to suggest the appropriateness of a behavioral choice but also as a tool to increase emotional arousal. For example, if the user performs an incorrect action, (s)he does not just read a canvas that signals the error but experiences the negative consequences firsthand. As reported by previous work in other emergency context, the inclusion of emotionally intense consequences of typical mistakes could promote knowledge retention (Chittaro & Buttussi, 2015) and therefore represents an example of a psychologically appropriate way to provide information.

In the following subsections, the salient points that emerged during the co-design activities and on which the "Safer Water" simulation was based, are reported in the form of guidelines. Although many of the guidelines are related to the specific situation of breakdown of the embankment, others may be generalizable to other emergency contexts.

2.2.6 Guidelines for virtual simulation design of flood emergency

This research provides information relating to the design of virtual simulation about flood emergency. The guidelines proposed below are the result of the salient findings that emerged from the co-design sessions. All the guidelines were organized into groups. In the first one, general information regarding the VE design related to flood emergency (scenarios and system design) is reported. The second group presents natural risk indicators of a flood emergency related to the possible breaking of an embankment (physical indicators of danger). Finally, the third one refers to which actions should be taken and which should be avoided during the flood emergency (facing the emergency).

Scenarios and system design

- The simulation should reproduce the salient event relating to the breakdown of the embankment.
- The scenario should include situations that allow the user to face the emergency by actively experiencing the effects of his/her behavior.
- The scenario should provide clear and immediate feedback on the correct and incorrect choices made by users.
- The scenario should include not only emergency management situations, but also risk identification.
- The scenario should include salient indicators of danger.
- In the simulation, the indicators of danger should be represented with realism and their timing.
- The simulation should include emotional stimuli to make the experience realistic.
- The simulation should include social stimuli that recall an ecological situation.
- Gamification elements should be included in the simulation to involve the user in the educational experience and not generate a situation of excessive stress for the user.
- The architecture of the scenario should be modular for overall or partial use of the simulation, according to the different objectives and contexts.
- The scenario should include environmental sounds related to flora and fauna of the river area to make the VE more realistic.
- The scenario should include the implementation of spatialized audio to increase the level of immersion.
- The scenario should include facilitators (e.g., bridges, hydrometers) to help users identify physical indicators of danger.
- Natural barriers should be inserted in the environment to limit ecologically the space that can be explored by the user.

Physical indicators of danger

- Presence of rising water from the subsoil (e.g., flooded manhole, puddles).
- Presence of high river level.
- Presence of muddy water.

- Presence of floating materials long the bed-river (e.g., branches, waste).
- Presence of whirling eddies along the bed-river, with particular reference to the riversides.
- Presence of leakage of cloudy water near the embankment (“fontanazzi”).

Facing the emergency

- If possible, prefer paved areas over the ground.
- In an outdoor situation, reach sites at a safe altitude and avoid areas of depression (e.g., avoid underpasses).
- In an indoor situation, go to the upper floors and avoid basements.
- Inside buildings, disconnect the power.
- Inside buildings, avoid using elevators.
- Avoid contact with electrical sources.
- Reach any designated collection points in case of municipal civil protection plans.
- Avoid saving movable objects from the lower floors to the upper floors (only basic necessities).
- Avoid crossing bridges when the water level is very high.
- Use the telephone lines only to contact emergency services.
- Do not drive in flooded areas.
- Do not drink tap water during an emergency situation.

2.2.7 Conclusions

The present work describes how co-design methodologies can be used to identify the contents of a VR simulation for river flood emergencies. Two affinity diagram sessions and a brainstorming were conducted with experts with different theoretical backgrounds, while keeping in mind that the ultimate goal was to improve the quality of life of citizens exposed to the risk of floods. From the activities, a series of results emerged. First, the adopted method made it possible to design a VR application that focused on two main aspects of a flood situation: the identification of the signs of a probable breakdown of the embankment and the adoption of the correct behaviors during the emergency. Second, the key points that emerged allowed us to highlight a set of guidelines to support the design of VR simulation for the purposes discussed in the work.

2.3 Evaluation of immersive VR experience

2.3.1 Background

The investigation of systems that engage users such as Head-Mounted Displays (HDM) is studied mainly from the perspectives of human–computer interaction. In this framework assume fundamental relevance the UX, considering that UX with interactive systems can be analyzed from multiple perspectives, giving the scope to the product and the interaction between user and product (Forlizzi & Battarbee, 2004). The VR experience is examined in particular by considering interaction techniques, relevant psychological aspects (e.g., presence and engagement), and side effects caused by the VR experience (Kim et al., 2020).

In terms of effectiveness, there are several aspects that need to be considered to design educational immersive VR experiences aimed at emergency preparedness. Mikropoulos and Natsis (2011) identified immersion and presence as key factors for fostering learning in VR environments. Immersion is a quantifiable characteristic of the VR system that refers to the objective level of sensory fidelity, while presence is the user's psychological response within the VR system (Slater et al., 1996; Slater, 2003). Immersive VR approaches can be more engaging than traditional teaching methods due to immersion and interaction. Particularly, immersion isolates the user from distractions, while interaction enables non-passive learning (Li et al., 2017).

Designing effective educational VR experiences for emergency preparedness requires considering two types of knowledge transmission: practical learning and skill acquisition, and conceptual learning (Mòrelot et al., 2021). For instance, knowledge of what is appropriate and inappropriate behavior during a flood is practical knowledge, while understanding flood alert indicators and the most at-risk areas is concept knowledge.

Chittaro and Butussi (2015) have highlighted how emotional activation in the context of immersive VR is linked to greater knowledge retention compared to the analogical mode of presentation of an emergency training. The authors were specifically inspired by entertainment games because emotional activation and involvement are typical features of these games. In order to maximize the effectiveness of the experience, some main techniques were extrapolated, such as the occurrence of unexpected events and vivid consequences when wrong behaviors take place. However, emotions do not always have a positive impact. Although they enhance perceptions of realism and knowledge retention, they have also been shown to deteriorate performance (Li et al., 2017).

Serious games (SG) in immersive VR environments can provide engagement, fun, instant feedback (Smith & Ericson, 2009), motivation for learning (Baboo et al., 2022), and high levels of interaction (Checa & Bustillo, 2020b). Therefore, to stimulate learning in various fields, it is possible to use serious games (SG) in immersive virtual reality environments that are defined as "*video games whose primary purpose is education, training, simulation, socialization, exploration, analysis, and warning rather than mere entertainment*" (Michael & Chen, 2006, p. 17). However, the use of immersive VR SG approaches has not always gained success in terms of learning outcomes. In fact, some researchers have hypothesized that in some cases the negative results may be due to a lack of experience with commands of the devices that allow the use of immersive virtual reality simulations (Checa & Bustillo, 2020a). A further hypothesis concerns that immersive VR could cause cognitive overload and distract the user (Makransky et al., 2019). Different factors impact cognitive load in IVR experiences: on the one hand there are individual differences and previous experience that cannot be controlled, on the other hand there is virtual simulation design related to environment and task aspects that can be improved by the researcher (Han et al., 2021). These hypotheses highlight the importance of learnability of the system interactions but also the need to investigate the effects of excessive cognitive load on learning outcomes in order to design an effective educational experience.

Another aspect that needs to be considered by the designer concerns the attitudes and the perceived preparedness of the user before and after this kind of immersive VR experiences. As highlighted by Rogers in the Protection Motivation Theory (PMT; 1975) it is important to effectively convey the danger of the events that occur during emergencies. According to the PMT theory, when the user perceives a threat, he makes two assessments, one linked to the danger of the event and one linked to the resources he has available to deal with it. If the individual considers the threat as not concrete enough, excessive or does not perceive that he has adequate resources to deal with it, it is likely that he implements negative/dysfunctional coping strategies, such as risk denial, avoidance, or defensive reactions. This is why it is very important to convey information that allows for the realization of an adequate perception of the danger in terms of attitude and perceived preparedness.

In summary, designing effective educational immersive VR experiences for flood emergencies requires considering multiple factors. Therefore, it is important to evaluate them with different stakeholders during the interaction design process.

2.3.2 Preliminary experts' evaluation and implementation of suggestions

Based on co-design activities, a prototype of an immersive VR experience aimed at supporting river emergencies has been developed. Before testing this prototype with target users, a preliminary evaluation with experts has been conducted in order to obtain useful feedback for the improvement of the simulation.

Two experts in the hydrogeological field were involved in the evaluation of the prototype to assess the realism of the simulation and gather feedback to enhance the experience. One of the experts (E1) was 42 years old at the moment of the evaluation, he obtained a Ph.D. in environmental hydrology and had over 10 years of expertise in the hydrogeological field. The other expert (E2) was 29 and graduated in civil engineering with a hydraulic address with more than 3 years of expertise in the hydrogeological field. Neither had prior experience with immersive VR technology or played video games frequently (less than once a year).

Before the immersive VR experience, a familiarization session with the technology was carried out, in which the commands to be used in the virtual simulation were briefly introduced. During the immersive VR experience, the experts individually offered observations and comments through the methodology of Thinking Aloud (Jääskeläinen, 2010). E1 experienced difficulties in understanding commands and interactions, indicating the need for a structured tutorial for teaching interactions. Nonetheless, the participant had a positive evaluation of the sensory fidelity, except for the fact that "*On the riverside the water should be higher if there is an imminent flood*". E2 had no difficulties interacting with the environment but suggested some graphic improvement: "*In reality, whirlpools are smaller and near the obstacles in the river*"; "*Fountains must be smaller in the center, with more water around it*".

The evaluation by the experts allowed for improving the realism of the prototype and designing a tutorial (Kao et al., 2021) a phase of the experience aimed at facilitating familiarization with immersive VR commands before the educational experience. The following section describes the final version of the simulation tested by users.

2.3.3 Description of immersive VR experience

The final version of the scenario features a stretch of river crossed by a bridge and flanked by a town. At a preliminary stage, in an area implemented ad hoc within the river scenario, the user performed a context-sensitivity tutorial. The context-sensitive tutorial phase of the experience had been designed in order to teach users the commands of the controller that allowed them to interact with the virtual environment. The tutorial consisted of four distinct steps, one for each of the four interactions that participants needed to learn: horizontal movement, vertical movement, object selection, and help functionality. Textual instructions anchored to the right-hand controller, namely tooltips, were provided to teach commands for each step. The first step focused on teleportation (this movement method has been selected because it is to move around the environment using the right-hand controller stick; Bozgeyikli et al., 2016). The second step taught vertical movement, which used the same command as horizontal movement but required pointing the controller upward. In the third step, participants learned how to select elements in the environment using ray casting and the right-hand controller trigger. The final step of the tutorial focused on the help modality command, which participants could use to recall all the commands anchored at the right-hand controller. The tutorial phase concluded with a test of learned commands: an audio recording prompted participants to implement each of the commands they had just learned when they reached the red-and-white barrier depicted in Figure 2.5.

Figure 2.5: End of the familiarization phase of the commands (tutorials) to be used during the immersive VR experience.

Then, the barrier opened and he/she started the flood emergency training. Since then, the user impersonated a member of the Civil Protection who had been called to carry out an inspection and understand if it was necessary to declare a state of emergency. The experience spanned two-time domains: one prior to the embankment break and the other following this event. The experience was divided into two phases, each characterized by a specific task. In the first phase of the experience, the user received a virtual phone call after he passed the red-and-white barrier that introduced the first task. In particular, through this call a colleague urged him/her to look around and report the presence of any danger indicators near the embankment within 5 minutes. As represented in Figure. 2.6, near the embankment, 10 target stimuli (i.e., danger indicators) belonging to 5 main categories were positioned.

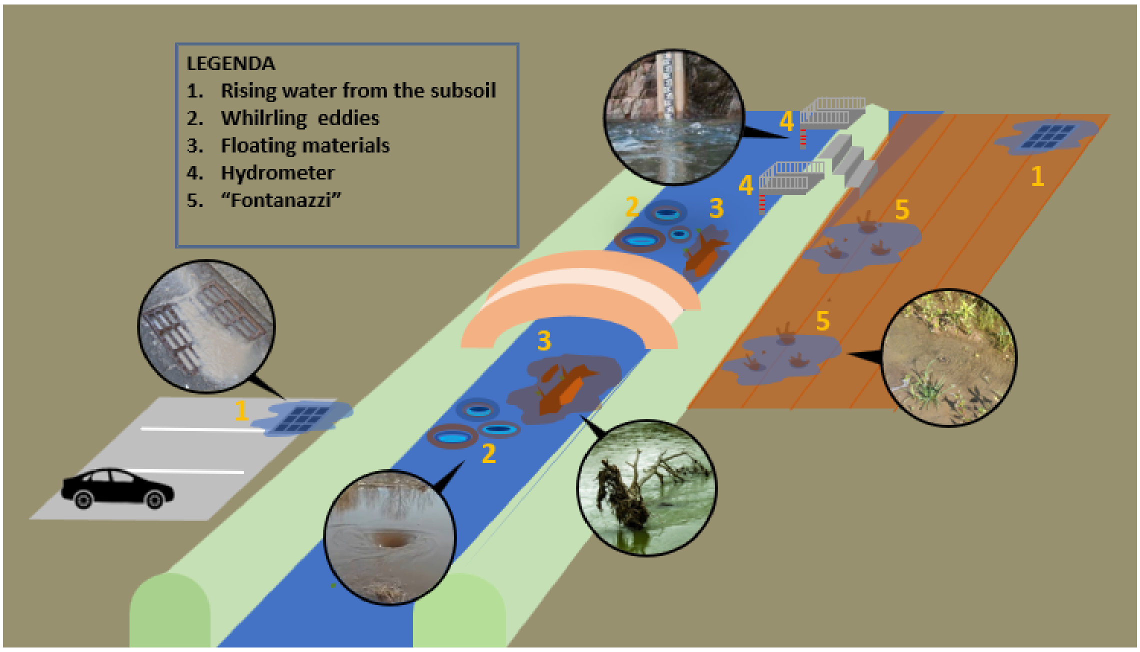

Figure 2.6: Graphical representation of the categories of danger indicators: rising water from the subsoil (e.g., flooded manhole, puddles), high river level and muddy water (that could be seen by the hydrometer), floating materials long the bed-river (e.g., branches, waste), whirling eddies along the bed-river, with particular reference to the riversides and finally, leakage of cloudy water near the embankment ("fontanazzi").

The user had a virtual tablet on which there is the list of the categories hidden by a series of "?". The name of each category was revealed as the participant selected the flood danger indicators located near the embankment. The virtual tablet not only allowed to view the category and the number of missing indicators as the user identified the indicators of the different categories, but also a timer that showed him/her the time available to complete the inspection (Figure 2.7).

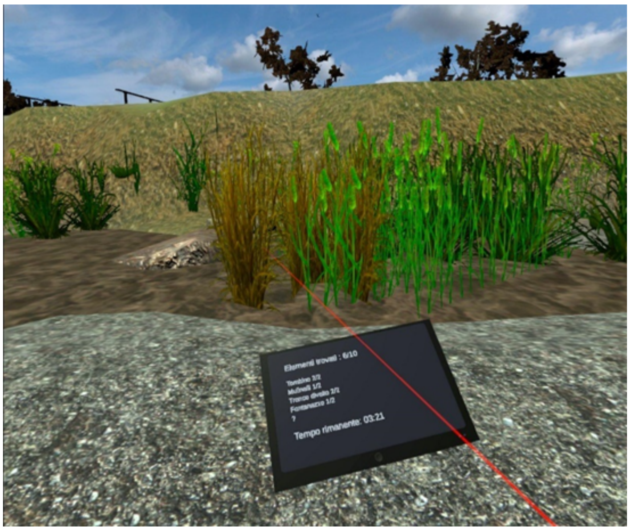

Figure 2.7: The user is looking at the virtual tablet where he can see the missing indicators to be found.

This exploratory phase ended when the user selected all of the danger indicators or at the end of the 5 minutes allocated to this phase. In both cases, the colleague started a phone call again by introducing a new task, that is to reach a position indicated on the virtual tablet where other colleagues from the Civil Protection would arrive to secure the area (Figure 2.8).

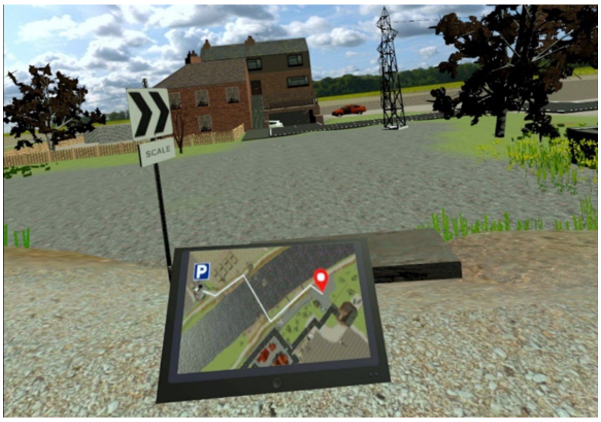

Figure 2.8: The user is following the indications on the map in the virtual tablet.

The point indicated on the map was located near the river, and once the participant reached it, the river bank broke, introducing the user to the final phase of the experience: escape from the flood to a safe place. This time the task was introduced by shouts in the distance urging them to save themselves. The user had to run away and choose a safe place; in the area there were different possibilities of success and failure as represented in Figure 2.9.

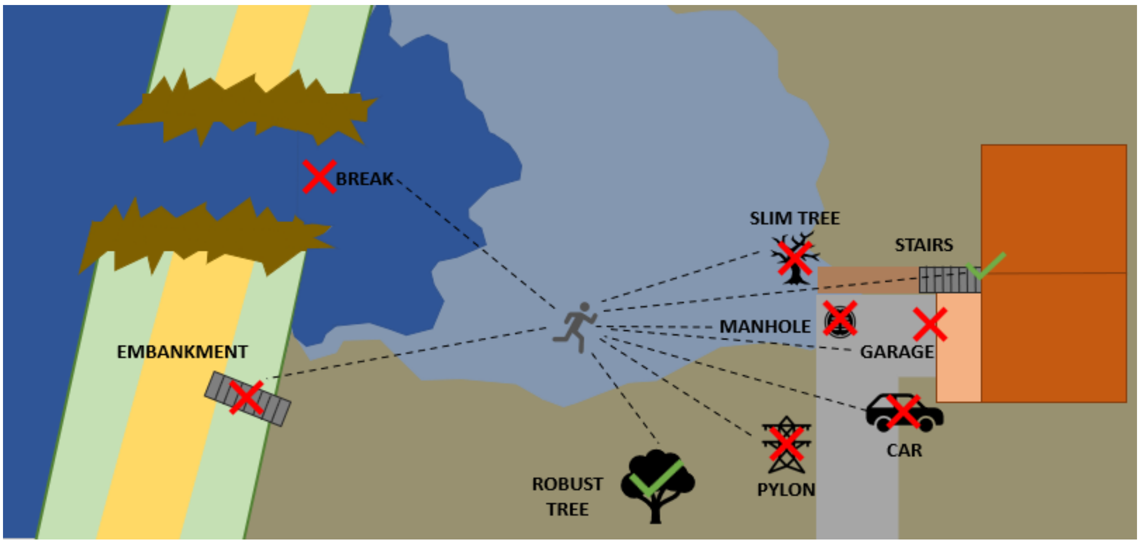

Figure 2.9: Graphical representation of success and failure possibilities in the escape phase.

Once the user had reached a safe place, call options appeared on the virtual tablet (Figure 2.10). In particular, the user could choose whether to contact family members or emergency services. The experience ended when the correct contact was selected, i.e., the emergency number (calling the family is inappropriate as it contributes to the clogging of telephone lines).

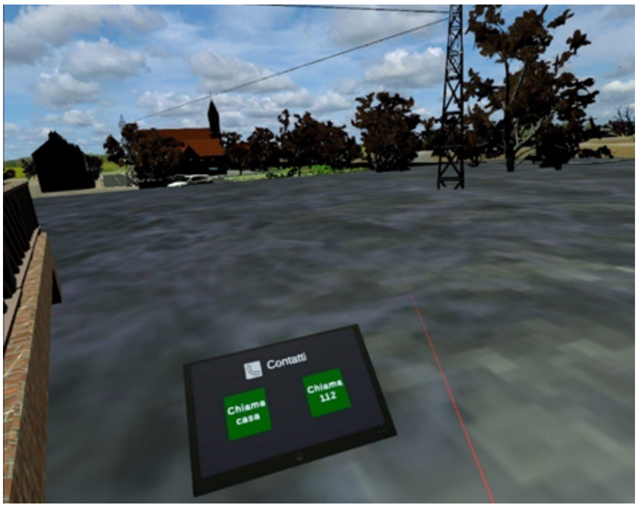

Figure 2.10: The user has to choose if calling the family number or the emergency one.

The experience concluded by providing the participant feedback on his/her performance, in particular the danger indicators found, the number of wrong choices in the escape phase, and the time taken to complete the experience.

2.3.4 Evaluation of immersive VR experience with target users

2.3.4.1 Materials and method

Different aspects of the experience were investigated and evaluated at different times and with different self-report materials and behavioral data.

Measure

Knowledge acquired. It is the acquisition of both conceptual and practical knowledge (Mòrelot et al. 2021), encompassing an understanding of flood alert indicators and high-risk areas, as well as appropriate behaviors to be adopted during floods. To assess the effectiveness of the immersive virtual reality experience in promoting actual learning, participants' knowledge was measured at three different time points: pre-IVR experience (T0), immediately post-IVR experience (T1), and 15 days post-IVR (T2).

Perceived preparedness and attitude towards flooding. Perceived emergency preparedness refers to an individual's perception of their ability to cope with a flood, while attitudes towards emergencies refers to how important participants considered the knowledge and practices transmitted through the experience and investigated people attitudes towards the possibility of being involved in a fluvial emergency. These dimensions are crucial for assessing the effectiveness of an immersive virtual reality experience, and were measured at T0, T1 and T2.

Learnability. It constitutes one of the usability dimensions according to Nielsen (1993) and can be considered as what determines whether a system allows a new user to efficiently use it in a short time. To measure this construct, log data was collected from the command test regarding the time and number of errors made. Self-reported learnability was also measured, and a maintained learnability test was performed to verify if participants remembered the command locations 15 days after the experiment.

Performance. User performance was measured based on the two phases that the experience was divided into. In the first phase, user performance was identified by the number of indicators identified, while in the second phase, it was identified by the number of attempts made before reaching a safe place and calling for help.

User experience (UX). It can be distinguished between perceived ergonomic quality, perceived hedonic quality, and perceived attractiveness of a product assuming that perceived ergonomic quality and perceived hedonic quality describe independent dimensions of the user experience. Therefore, it assumes that users perceive several distinct aspects when they evaluate an interface. Ergonomic quality focuses on the goal oriented or task-oriented aspects of product design, while hedonic quality focuses on the non-task-oriented quality aspects of the interface. The perceived attractiveness of the product is then a result of an averaging process from the perceived quality of the software concerning the relevant aspects in a given usage scenario (Hassenzahl, 2001; Laugwitz & Schrepp, 2008).

Emotional engagement. It refers to both positive and negative affects felt by the user during the experience. Positive affects reflect the level of pleasant engagement, the extent to which a person feels enthusiastic, excited, active, and determined; while negative affects reflect a general dimension of unpleasant engagement and subjective distress that subsumes a broad range of aversive affects including fear, nervousness, guilt, and shame (Watson, 1988).

Sense of presence. As anticipated, presence can be considered the subjective result of immersion. This construct is characterized by four dimensions: engagement, sensory fidelity, adaptation/immersion, and interface quality. Engagement follows the focusing of one's mental energy and attention on a coherent set of stimuli or activities/events that are significantly correlated. Sensory fidelity is the degree to which a virtual reality environment allows users to examine objects both visually and through touch, including auditory aspects. Adaptation/immersion, in this case, was considered as the perception of feeling wrapped up and included in an environment that provides a continuous flow of stimuli and experiences to interact with. Finally, interface quality is the degree to which control or display devices interfere or distract from activity performance, but also the extent to which participants felt able to concentrate on tasks (Witmer & Singer, 2005).

Cognitive load (CL). It can be understood as the amount of working memory used in performing a particular task and can be divided into intrinsic, extraneous, and germane. Intrinsic load depends on the difficulty and complexity of the material being learned, extraneous load arises from extraneous processes that do not directly contribute to learning, often fueled by inadequate instruction design and the presence of distractions, and germane load is relevant to the material being learned (Sweller, 2011; Albus et al., 2021).

Self-report data

Knowledge on flood emergencies questionnaire. Conceptual and practical learning was measured using a questionnaire developed ad hoc that included two questions on conceptual learning (e.g., "what are the main indicators of flood warning status?") and five questions on practical learning (e.g., "what are the correct behaviors to adopt during an indoor disaster flood?"). The scores were calculated based on a scoring scheme approved by two independent judges. Scores were calculated based on a scoring scheme agreed upon beforehand by two independent judges, totaling 30 points (9/30 for conceptual knowledge and 21/30 for practical knowledge).

Perceived preparedness and attitude towards flooding questionnaire. To measure perceived preparedness for floods, three items related to their knowledge and practical competence related to floods and fluvial emergencies were created based on the "familiarity with activities and terms" dimension from the Emergency Preparedness Information Questionnaire by Wisniewski and colleagues (2004). Specifically, participants were asked to rate their perceived capability to detect flood warnings and their perceived ability to implement appropriate behavior during a fluvial emergency. To measure attitudes towards emergencies, two items were constructed based on the "attitudes towards disaster and emergency preparedness" dimension by Nofal et al. (2018). Participants responded through a 7-points likert scale from "Completely disagree" to "Completely agree").

Learnability questionnaire. As in Kao et al. (2021), the learnability of the commands was measured through the command subscale of the Player Experience of Need Satisfaction (PENS) (Ryan, Rigby & Przybylski, 2006) consisting of 3 items to which participants had to respond by reporting their degree of agreement using a 7-point Likert scale where 1 indicated "Completely disagree" and 7 indicated "Completely agree". One of the proposed items is reported below: "Learning the commands to use in the experience was easy." Two additional ad hoc items were also integrated to examine the user's perception regarding the speed of learning the commands, as the less time it took to learn the commands, the greater the learnability, and their need to ask for help to understand how the commands worked since the need to ask for external help would be an indicator of an interface that does not effectively assist the user (Linja-aho, 2006).

Lernability retention questionnaire. An ad hoc questionnaire was created to measure how well the participants remembered the commands used during the experience 15 days after completing it. Multiple-choice answers asked them to associate the command with its corresponding function. To avoid the facilitation effect due to the fact that the interactions to be remembered were only 3 (movement, selection, "command map"), items related to commands not used during the experience were also included in the questionnaire, to which the response "None of the above" had to be attributed, and an interaction not present during the experience was added, namely the possibility of skipping.

User experience questionnaire. It consists of 26 items designed by (Hassenzahl, 2001) and is composed of seven dimensions. Attractiveness (6 items), refers to the general impression of the IVR experience. Pragmatic qualities are related to the tasks the user has to achieve while using the product and describe its efficiency (4 items), perspicuity (4 items), and dependability (4 items). Hedonic qualities describe how stimulating a product is (4 items) and how it gives a sense of innovation (novelty; 4 items). The items have the form of a semantic differential, i.e. each item is represented by two terms with opposite meanings placed at the extremes of a 7-point scale.

PANAS questionnaire. To measure emotional engagement, the Positive and Negative Affect Schedule (PANAS) by Watson and colleagues (1988), translated by Terraciano, McCrae, and Costa (2003) was used. The self-report questionnaire is reliable and valid for the assessment of positive and negative affects. The questionnaire consists of ten items on positive affects and ten items of negative affects (e.g., interested, strong, scared, nervous) which can be answered using an ordinal scale between 1 and 5, where 1 stands for "very little/not at all" and 5 stands for "extremely".

PQ questionnaire. An adaptation of the Presence Questionnaire (Witmer et al., 2005) composed of 27 items was used to measure sense of presence. In the version we used, the items relating to haptic feedback were also removed (items 13, 17) as this sensory modality was not relevant to the IVR experience.

Cognitive load questionnaire. Cognitive load was measured through an adapted version of the Multidimensional Cognitive Load Scale for Virtual Environments (MCLSVE) by Andersen and Makransky (2021). The scale validated by the authors aims to measure CL through the distinctions highlighted by Sweller in Cognitive Load Theory (Sweller, 2011), which are the same as those of Albus, Vogt, and Seufert (2021). The tool used to measure cognitive load was adapted based on the immersive virtual reality simulation of the emergency theme experienced by the user. The response scale used was a Likert scale ranging from 1 to 7 points, where 1 indicated "completely disagree" and 7 indicated "completely agree."

Log data

From a quantitative standpoint, the users' behaviors during the experience were recorded through a tracking script embedded in the program that constitutes the experience. The user's position within the virtual environment was saved through coordinates in the three main axes (x, y, z), and the direction in which the user was looking was recorded through rotation on these three axes. The script allowed for an interval of 0.1 seconds for each position update (i.e., a frequency of 10 updates per second, with possible slight fluctuations due to the processing power of the headset). The program code divided the experience into 5 phases: tutorial, searching for danger indicators (which ends with the discovery of all indicators or after 5 minutes), moving to the breach zone, breaking through the barrier and escaping, and the final phase (including the moment of the call - to home or emergency services - and the conclusion of the experience). The tracking script recorded the exact moment when each phase began and recorded their main events, the moment they occurred (based on seconds elapsed from the start of the experience), and the user's position at that time. Specifically, during the tutorial, the collected data related to the execution times of the different sections and interactions performed during this phase, during the search for indicators phase, the discovery of each individual indicator was recorded (i.e., when it was selected), and finally, during the barrier-breaking and escape phase, the type of target chosen by the user to reach safety was recorded when the user reaches it. If an incorrect target is chosen, and a new attempt is made, the moment the simulation restarts from the barrier-breaking is recorded; if the correct target is chosen, the moment when the user calls home or emergency services is recorded. In addition, the execution times of the command test were recorded, which allowed us to highlight the effectiveness of the different types of tutorials once they were completed.

Procedure

On the evaluation day, each participant was given an information note and an informed consent, which had to be read, understood, and signed before proceeding with the experimental session. The experimental session included the filling of a socio-demographic questionnaire and a questionnaire concerning the assessment of basic knowledge in the field of flooding. Then, the participant was shown a video through the 360 viewer that introduced the material on fluvial emergencies necessary to carry out the IVR experience. Afterwards, a familiarization phase where participants were given a tutorial aimed at teaching the system control that would later be used to interact in the IVR SG in an emergency situation. Then, the participant carried out a short test in the IVR where he/she would be asked to perform the interactions he/she had just learnt. Subsequently, the designed immersive virtual reality experience began. In the first phase, the participant had to search for danger indicators (a total of 10) that might presage a possible flood in the riverbank's area. In the second phase, which followed the breach of the riverbank, the participant had to reach a safe place as quickly as possible and search for help (Figure 2.11).

Figure 2.11: A participant experiencing the virtual simulation.

At the end of the immersive virtual reality experience, the participant filled out a series of questionnaires assessing the learnability of the controls, knowledge learned, user experience, emotional engagement, sense of presence, and cognitive load. All questionnaires were administered through the use of the Qualtrics ^{XM} platform[4]. The total duration of the study was about one hour.

The present evaluation was designed in accordance with the Declaration of Helsinki and approved by the ethics committee for psychological research at the University of Padova.

The steps described are synthesized in the figure below (Figure 2.12).

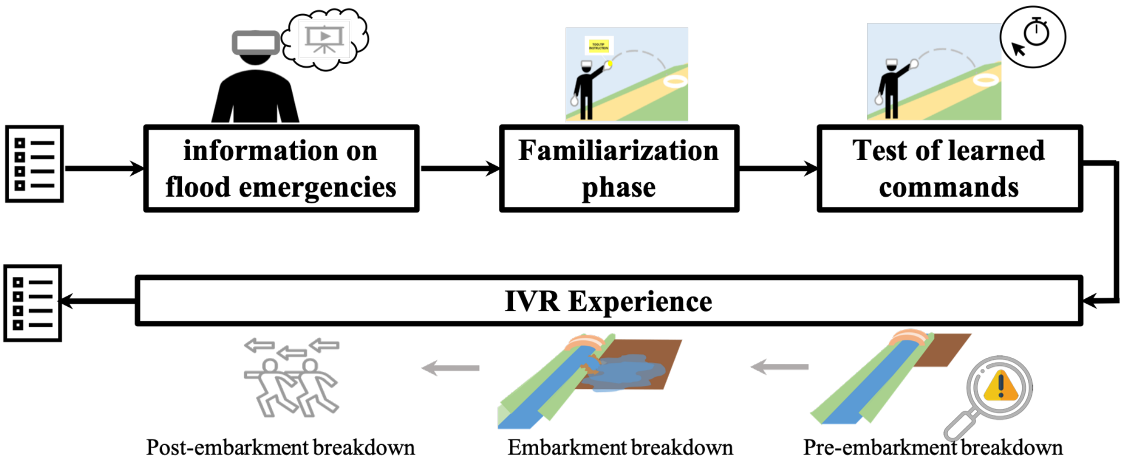

Figure 2.12: Graphic representation of the procedure steps.

2.3.4.2 Participants

The study included 31 participants (F =15), with an age range between 18 and 30 years (M age= 24.39, SD = 3.03). The recruitment criteria required participants to be right-handed, as the virtual environment used in the experiment was designed for right-handed users. The age range was also limited to 18-30 years (emergent adulthood; Arnett, 2000) to control the effect of cohort affiliation on the collected data. In addition, participants were not required to have experienced a flood, worked in civil defense or fire department, or participated in training events related to river emergencies. Furthermore, participants with prior experience in flood-related emergencies were excluded due to a factor of sensitivity, or were frequent users of head-mounted displays (HMDs) to assess the learnability of the experience commands. This criterion was important to ensure that the virtual experience was usable and effective even for first-time users.

[4] https://www.qualtrics.com

Particularly, none of the participants reported using Virtual Reality daily or weekly. Some (3/31) use it about once a month or have tried it only once (8/31). Most (20/31) have never used VR technologies. Frequency of video game use shows a similar pattern to VR use, with most participants reporting that they never play video games (14/31) or play about once a month (10/31), while only a small portion (7/31) play more than once a week. The sample has an average schooling of 15.16 years (Min = 8, Max = 18, SD = 2.58). More than half of the sample (17/31) are students, while the remainder are employees (7/31), teachers/researchers (3/31), doctors in training (3/31) and blue-collar workers (1/31).

2.3.5 Analysis and Results

The following are the analyses and results of the self-report and behavioral data collected. Analysis was performed using R study software. A series of Shapiro-Wilk tests were conducted to check whether the assumption of normality in the data was met. The tests showed a non-normal distribution, so nonparametric analyses were performed. The results are shown in the table below.

2.3.5.1 Knowledge of flood emergencies

In the following paragraphs knowledge results collected by self-report questionnaire are presented. In particular, the percentage scores of prior knowledge recorded before the immersive VR experience (T0), the knowledge acquired after the end of the immersive VR experience (T1) and the retention of knowledge recorded 15 days after the end of the experience (T2) were calculated. Scores were calculated for conceptual knowledge (maximum score 9/9), practical knowledge (maximum score 21/21) and total knowledge (maximum score 30/30).

From a percentage point of view, the results highlighted that in T0 both in practical knowledge and in conceptual knowledge participants obtained less than 20% of the maximum score. In particular, they gained an overall total knowledge of 14.41% (14.13% in practical knowledge and 15.05% in conceptual knowledge).

In T1 the knowledge scores obtained had incremented, in fact, both knowledge types were over 40% of the maximum score. In particular, they gained a total knowledge of 52.04% (46.24% in practical knowledge and 65.59% in conceptual knowledge).

Finally, the knowledge scores obtained in T2 were again over 40% of the maximum score, but lower than in T1. In particular, they gained a total knowledge score of 47.85% (41.78% in practical knowledge and 62.01% in conceptual knowledge; Figure 2.13).

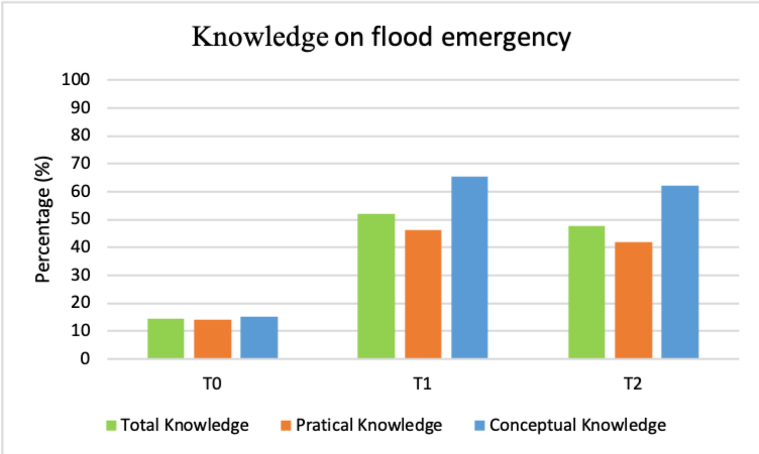

Figure 2.13: Knowledge percentage scores (%) obtained by the participants both total and considering separately conceptual and practical learning in the 3 moments of measurement (T0, T1, T2).

Subsequently, it was evaluated how much the users' knowledge of flood emergencies improved once they have done IVR experience compared to the prior knowledge (T1-T0), how much the knowledge has been maintained 15 days after the experience (T2-T1), and the effective knowledge acquired (T2-T0).

Descriptive data on percentages highlighted that from T0 to T1 participants increased their total knowledge score by a percentage of 37.63% (32.10% in practical knowledge and 50.54% in conceptual knowledge). Between T2 and T1 participants decreased their total knowledge score by a percentage of 4.19% (4.45% in practical knowledge and 3.58% in conceptual knowledge). Finally, considering the effective knowledge acquired from T0 and T2, the total knowledge score increased by a percentage of 33.44% (27.65% in practical knowledge and 46.95% in conceptual knowledge; Figure 2.14).

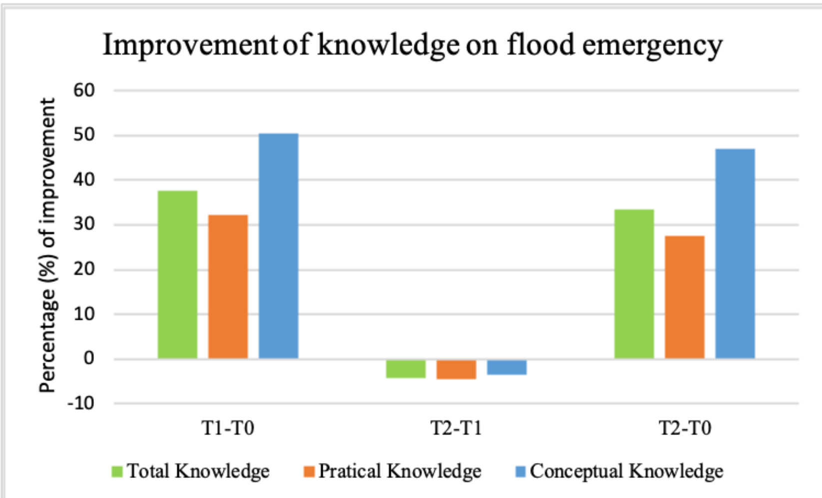

Figure 2.14: Differences between knowledge percentage scores (%) obtained by the participants both total and considering separately conceptual and practical learning in the 3 moments of measurement (T0, T1, T2).

To test whether immersive virtual reality training was effective, a series of Friedman tests (B.H. corrections) were conducted to compare participants' a priori and acquired knowledge and retention. The results are shown in the table below (Table 2.3).

Knowledge of flood emergencies						
Dimensions	Time	Descriptive statistics			Friedman Test	
		Mdn	M	SD	$X^2(2)$	p (B.H.)
Total Knowledge ***	T0	4.00	4.32	2.39	48.46	< .001
	T1	16.00	15.61	3.72		
	T2	14.00	14.36	2.95		
Conceptual Knowledge ***	T0	1.00	1.35	0.88	53.09	< .001
	T1	6.00	5.90	1.30		
	T2	5.00	5.58	1.09		
Practical Knowledge ***	T0	3.00	2.97	1.80	46.12	< .001
	T1	10.00	9.71	2.77		
	T2	8.00	8.77	2.20		

*Table 2.3: Results of Friedman tests which compares the knowledge of flood emergencies in the three different times T0, T1, and T2; medians, means and standard deviations of the self-report questionnaire concerning flood emergency knowledge (*p(B.H.) < .05).*

A series of post-hoc tests were performed through Wilcoxon's tests (B.H. corrections). The tests revealed that the overall knowledge about floods was significantly increased (z = -4.86, p < .001) from T0 (Mdn = 4.00, M = 4.32, SD = 2.39) to T1 (Mdn = 16.00, M = 15.61, SD = 3.72) In addition, statistically significant differences emerged (z = 2.74, p < .01) comparing T1 (Mdn = 16.00, M = 15.61, SD = 3.72) and T2 (Mdn = 14.00, M = 14.36, SD = 2.95), showing how not all the knowledge acquired has actually been maintained over time. Furthermore, there was a maintenance of knowledge (z = -4.84, p < .001) from T0 to T2 (Mdn = 14.00, M = 14.36, SD = 2.95). Similarly, for the practical knowledge, tests showed a statistical increase at T1 (z = -4.78, p < .001; Mdn = 10.00, M = 9.71, SD = 2.77) and T2 (z = -4.83, p < .001; Mdn = 8.00, M = 8.77, SD = 2.20) respect to T0 (Mdn = 3.00, M = 2.97, SD = 1.80). In addition, statistically significant differences emerged (z = 2.28, p = .02) comparing T1 and T2, showing how not all the practical knowledge acquired has actually been maintained over time. With regard to conceptual knowledge, it was found that was significantly increased (z = -4.86, p < .001) at T1 (Mdn = 6.00, M = 5.90, SD = 1.30) with respect to T0 (Mdn = 1.00, M = 1.35, SD = .88) and at T2 (z = -4.86, p < .001; Mdn = 5.00, M = 5.58, SD = 1.09) respect to T0. Finally, there were no statistically significant differences when comparing T1 and T2 (z = 1.72, p = .8).

2.3.5.2 Perceived preparedness and attitudes towards flood emergencies

A series of Friedman tests were conducted to test for any change with respect to users' perceived preparedness and their attitudes toward emergency situations. The results are shown in the table below (Table 2.4).

Perceived preparation and attitude towards flood emergency						
Dimensions	Time	Descriptive statistics			Friedman Test	
		Mdn	M	SD	X^2	p (B.H.)
Attitudes toward Flood Emergency ***	T0	2.50	3.03	1.51		
	T1	5.50	5.45	1.08	30.50	< .001
	T2	5.50	5.21	1.00		
Perceived Preparedness ***	T0	2.66	2.80	0.88		
	T1	5.67	5.66	0.60	50.60	< .001
	T2	6.00	5.60	0.57		

*Table 2.4: Results of Friedman tests on users' perceived preparedness and their attitudes toward emergency situations; medians, means and standard deviations of the self-report questionnaire concerning perceived preparedness and attitudes towards flood emergencies (***p(B.H.) < .001).*

A series of post-hoc Wilcoxon tests (B.H. correction) showed that attitudes toward the flood emergency were higher (V = 393.00, p < .001) post-training (Mdn = 5.50, M = 5.45, SD = 1.08) than pre-training (Mdn = 2.50, M = 3.03, SD = 1.51). Moreover, scores also remain higher (V = 423, p < .001) even 15 days after the experience (Mdn = 5.50, M = 5.45, SD = 1.00). No significant differences emerged with respect to attitude toward flood emergency post-training and after 15 days (V = -128, p = .06).

Similarly, perceived preparedness is increased (V = 496, p < .001) post-training (Mdn = 5.67, M = 5.66, SD = .60) than pre-training (Mdn = 2.66, M = 2.80, SD = .88) and is also higher (V = 496, p < .001) 15 days later (Mdn = 6.00, M = 5.60, SD = .57). No statistically significant differences emerged between perceived preparedness post-training and 15 days later (V = -14, p = .82).

2.3.5.3 Commands of immersive VR experience: subjective and objective learnability

In order to investigate whether the level of command's learnability was positive for the participants, a one-sample Wilcoxon test was performed by comparing the scores obtained by the

participants with the median of the scale (Mdn = 7). Results showed that users reported high scores in relation to learnability (V = 493.00, $p < .001$; Mdn = 5.57, M = 5.47, SD = .82.
Moreover, descriptive analyses related to the test of learned commands were conducted. The total median time that participants took to complete the total commands test was 3.17 seconds (M = 4.83, SD = 4.08). In particular, in the selection task, they took a median time of 1.79 seconds (M = 2.89, SD = 3.75), in the movement task of 0.82 seconds (M = 1.09, SD = 1.01) and in the help command of .26 seconds (M = 1.74, SD = 1.66). In addition, the total median error made by users was < 1 (Mdn = .00, M= .52, SD = .68); in the selection task they commit in mean .32 errors (SD = .65), in the movement task of 0.03 (SD = .18) and in the help command of .16 (SD = .37). Finally, another measure of learnability can be inferred by the frequency of use of the help modality. The function has been used only 1 time by 4 participants during the experience.

2.3.5.4 Performance in the immersive VR experience

As regards the first phase of the IVR experience, namely search of danger indicators, behavioral data relating to the number of target selected were collected (maximum score 10/10). Differently for the second phase, namely flood escape, the incorrect behaviors put in place by the participants before reaching safety were identified.
During the search of danger indicators phase of the experience, the users found on average about 8/10 of the danger indicators (Mdn = 8.00, M = 8.03, SD = 1.62). Specifically, 6 participants identified all the danger indicators, using an average time of 197.72 seconds (SD = 50.51) while 25 participants, not having identified all the danger indicators, took the 5 minutes foreseen by the experience.
During the flood escape phase of the experience, users took an average of 62.75 seconds to get to safety (SD = 30.45) and identified an average of 1.61 wrong targets (SD = 1.52). In particular, 8 participants completed the task by reaching safety without having to repeat the experience (Table 2.5).

Behavioral data of immersive VR experience			
	Mdn	M	SD
Danger indicators selected	8.00	8.03	1.62
Wrong target	1.00	1.61	1.52

Table 2.5: Medians, means and standard deviations of the danger indicators selected during the first phase of immersive VR experience and wrong safe place achieved during the second phase of immersive VR experience.

2.3.5.5 Users' opinions about the immersive VR experience

The followings are the analyses and results of the different self-report questionnaires that the user filled out to evaluate the experience. Negative items in the questionnaires were reversed before conducting the statistical analyses.

User experience. In order to evaluate the user experience with immersive virtual training, a series of one-sample Wilcoxon tests were performed by comparing the scores obtained by the participants with the median of the scale (Mdn = 4; from -3 to +3). The table below highlights how participants rated the experience positively in terms of attractiveness, perspiquity, novelty, stimulation, dependability, and efficiency (Table 2.6).

User Experience Questionnaire					
Dimension	Descriptive statistics			Wilcoxon Test – one sample	
	Mdn	M	SD	V	p (B.H.)
Attractiveness***	2.66	2.41	0.60	496.00	< .001
Perspiquity***	2.50	2.24	0.79	465.00	< .001
Novelty***	2.50	2.31	0.70	496.00	< .001
Stimulation***	2.33	2.27	0.73	496.00	< .001
Dependability***	1.25	1.14	0.74	455.00	< .001
Efficiency***	2.00	1.94	0.77	496.00	< .001

*Table 2.6: Results of one-sample Wilcoxon tests on user experience; medians, means and standard deviations of the self-report user experience questionnaire (*** p(B.H.) < .001).*

Emotional engagement. The emotional aspects that characterized the experience were identified by comparing the scores obtained by the participants with the median of the scale (Mnd = 3; from 1 to 5). In particular, as shown in the table below, it emerged that most of the emotions experienced by the participants concerned positive affects (V = 428.40, p < .001; Mdn = 3.90, M = 3.84, SD = .57), while most of the negative affects were statistically lower than the median of the scale (V = 1.1, p < .001; Mdn = 1.40, M = 1.57, SD = .55), thus resulting marginal (Table 2.7).

PANAS Questionnaire						
Affects	Emotion	Descriptive statistics			Wilcoxon Test - one sample	
		Mdn	M	SD	**V**	**p (B.H)**
Negative Affects	Distressed ***	2.00	1.94	0.89	16.00	< .001
	Upset ***	1.00	1.55	0.85	12.00	< .001
	Guilty ***	1.00	1.39	0.72	4.50	< .001
	Scared ***	1.00	1.55	0.89	10.00	< .001
	Hostile ***	1.00	1.06	0.25	< .001	< .001
	Irritable ***	1.00	1.26	0.51	< .001	< .001
	Ashamed ***	1.00	1.39	0.76	3.50	< .001
	Nervous ***	2.00	1.94	1.06	15.00	< .001
	Jittery ***	2.00	2.16	1.16	50.50	< .001
	Afraid ***	1.00	1.48	0.89	8.00	< .001
	Total	1.40	1.57	.55	1.1	< .001
Positive Affects	Interested ***	4.00	4.42	.62	435.00	< .001
	Excited*	4.00	3.61	1.05	189.00	.01
	Strong	3.00	3.00	1.18	85.50	1.00
	Enthusiastic *	4.00	3.61	1.15	236.00	.01
	Proud*	3.00	2.48	1.18	43.00	.02
	Concentrating ***	4.00	4.13	.85	383.00	< .001
	Inspired ***	4.00	4.32	.65	406.00	< .001
	Determined ***	4.00	3.94	.81	268.00	< .001
	Active ***	4.00	4.39	.67	406.00	< .001
	Attentive ***	5.00	4.45	.68	488.00	< .001
	Total	3.90	3.84	.57	428.50	< .001

*Table 2.7: Results of one-sample Wilcoxon tests on positive and negative affects during immersive VR experience; medians, means and standard deviations of the self-report PANAS questionnaire (***p(B.H.) < .001; *p(B.H.) < .05).*

Sense of presence. To evaluate the sense of presence relating to the immersive virtual reality experience, Wilcoxon tests (B.H. correction) were carried out by comparing the scores obtained by the participants with the median of the scale (Mdn = 4; from 1 to 7). The results highlighted that the participants felt involved and immersed in the experience that had just ended. They also reported a high sensory fidelity and interface quality of the simulation (Table 2.8).

Presence Questionnaire					
Dimension	Descriptive statistics			Wilcoxon Test - one sample	
	Mdn	M	SD	V	p (B.H.)
Involvement ***	4.64	4.74	0.83	421	< .001
Sensory Fidelity***	5.80	5.61	0.92	430	< .001
Adaptation/Immersion ***	5.38	5.46	0.76	465	< .001
Interface Quality ***	5.67	5.39	1.04	397	< .001

*Table 2.8: Results of one-sample Wilcoxon tests on presence during immersive VR experience; medians, means and standard deviations of the presence questionnaire (*** p(B.H.) < .001).*

Cognitive load. Finally, the participants evaluated the cognitive load required during the IVR experience, differentiating the intrinsic, extrinsic, and germane cognitive load. Wilcoxon tests (B.H. correction) were performed to compare the scores obtained by the participants with the median of the response scale (Mdn = 4; from 1 to 7). The results shown in the table below highlighted a low intrinsic and external cognitive load, otherwise, the German cognitive load was high (Table 2.9).

Cognitive load Questionnaire					
Dimensions	Descriptive statistics			Wilcoxon Test - one sample	
	Mdn	M	SD	**V**	**p (B.H)**
Intrinsic Cognitive Load ***	3.00	3.05	1.17	496.00	< .001
Extraneous Cognitive Load ***	1.83	1.80	.66	.00	< .001
Germane Cognitive Load ***	6.25	6.35	.60	43.00	< .001

*Table 2.9: Results of one-sample Wilcoxon tests on cognitive load during immersive VR experience; medians, means and standard deviations of the self-report cognitive load questionnaire (***p(B.H.) < .001).*

2.3.5.6 User preferences about training in flood emergencies

Finally, participants were asked through two purpose-made items whether they would have preferred to conduct training other than the proposed training. Wilcoxon's tests showed that participants would not have preferred face-to-face training (V = 3504.50, p < .001; Mdn = 6.00, M = 5.90, SD = 1.16) nor online training through the presentation of traditional materials (e.g., videos), (V = 3322.00, p < .001; Mdn = 6.00, M = 5.76, SD = 1.23).

2.3.6 Discussion

The purpose of the present study was to evaluate the IVR simulation "Safer Water" for river flood emergencies. The focus was on understanding whether it was an effective training and validating the guidelines that emerged from the previous co-design activities (Gamberini et. al, 2021). To do this, several stakeholders were involved. In the first evaluation phase, experts in the emergency field were involved to evaluate the implemented simulation, focusing on the educational contents and realism of the river environment. After the implementation of experts' suggestions and finalization of the simulation, the final version was tested with target users.

Results highlighted that there is an improvement in knowledge of flood emergencies both at the end of the immersive VR experience and 15 days later. Despite the knowledge learnt had not been all maintained, it is important to notice the efficacy of the experience. In fact, the effective knowledge improved is 27.65% for practical knowledge and 46.95% for conceptual knowledge.

This data shows that the simulation designed on the basis of co-design activities with experts works in terms of knowledge transfer to the user. The results also sustain other studies in the literature that have demonstrated the potential of IVR for educational purposes (Feng et al., 2018, Checa & Bustillo, 2020).

Furthermore, the experience helped the users gaining awareness of the flood risk and gaining adequate perceived preparation against this kind of emergencies, both important aspects to effectively convey the danger of the events that occur during emergencies (Chittaro et al., 2014; Rogers, 1975). In this perspective, virtual reality can be seen as a persuasive medium from this point of view that can impact on the user's attitudes and perceived preparation through some persuasive principles. The principle of cause and effect may have persuaded people to change their attitudes because it has allowed them to immediately observe the link between cause and effect, for example in the escape phase of the experience when they had to escape from the flood and choose a wrong target as a safe place (Fogg, 2002).

In order to teach commands to the users and familiarize with the environment, an immersive tutorial was developed at the beginning of the simulation. The tutorial has demonstrated its efficacy in teaching how to interact with the virtual environment, in fact the average time for carrying out the command tests is less than 5 seconds and the mean total errors was less than 1 ($M = .52$, $SD = .68$). These results have been sustained also by opinions of the users, that reported the ease and speed of learning commands. Positive results have already been found by other researchers on the familiarization modality used in our experience (Frommel et. al, 2017; Andersen, 2012).

Regarding practical qualities of the IVR experience, users reported to solve tasks with no unnecessary effort, felt in control of the interactions with the environment and found it easy to become familiar with the interaction modality.
The behavioral data highlighted how the users were able to identify an alert situation, in fact they identified most of the danger indicators (on average 8/10, SD = 1.62). As far as the second phase is concerned, however, the users did not always save themselves on the first attempt (average of errors 1.61, SD = 1.52). This may be due to the fact that the situation required a different emotional involvement due to the surprising event of the embankment break from which the users have to escape as fast as they can (Chittaro & Buttussi, 2015).
Participants declared to be emotionally engaged during the whole experience, mainly positively. In fact, they reported to be attentive, interested, determined, concentrated and inspired during the simulation. Negative affects were reported at least in a slight way. This result confirms the fact that the experience has been designed in order to engage emotionally the participants without creating an uncomfortable or traumatic experience. Through this aspect, the simulation can be accessed by a wide range of users and the impact of negative emotions is minimized (Li et al., 2017). Even the evaluation of hedonic qualities that have been reported in the UX questionnaire showed how the simulation has been experienced positively, in fact participants found it innovative, creative, interesting, and motivating.
Additionally, from users' opinions emerged high levels of sense of presence in terms of sensory fidelity, and interface quality, but also realism and immersion have been reported. Moreover, the participants did not consider the experience contents and procedures excessively complicated, they felt that the simulation improved their comprehension and knowledge about flood emergencies and reported low levels of extraneous load due to the instructions, the interactions and the environment have been found. These results sustain the importance of the design implementation of educational experiences in order to increase presence, decrease extraneous cognitive load and maximize knowledge outcomes (Schrader and Bastiens, 2012), constructs that seem to impact on learning outcomes (Huang et al., 2019).
Overall, it emerges that users were satisfied with the training experience related to flood emergency management, reporting that they prefer this type of immersive training compared to face-to-face trainings and courses offered online with traditional methods.

2.3.7 Conclusion

The present work describes the evaluation of VR simulation for river flood emergencies "Safer Water", an immersive virtual reality experience aimed at improving the quality of life of citizens that living near rivers or in areas prone to flooding. The immersive simulation proposed is realized through an interactive design process which has made it possible to obtain effective training to support river emergencies. The final version of the immersive training provides useful conceptual and practical knowledges to users on the event of an embarkment breakdown with subsequent flooding. Indeed, it allows increasing people's awareness and preparedness. The user's satisfaction and their preferences to use immersive VR training highlights the potential of these systems in the emergency context of flooding. So, the guidelines made in the early stages of design have been validated. The effectiveness of training channeled towards user engagement, as well as the relative interest in conducting such training with immersive virtual reality rather than traditional methods, shows the potential of immersive environments. IVR thus offers the opportunity to conduct effective courses both remotely or during face-to-face courses offering practical experience.

Some limitations of the present work should be acknowledged. The first one includes not having adequately considered how to manage flood emergency situations in indoor contexts. In fact, although in the affinity diagram sessions elements relating to the behaviors to be kept within domestic environments emerged, it was preferred in the resulting scenarios to focus on the outdoor context. Future studies will have to investigate this aspect, designing virtual simulations that help citizens adopt safe behaviors during a flood emergency, even inside their homes.

The users involved were citizens who live in areas at risk of young age. While this allowed for consistent data, it also limits knowledge with respect to the effectiveness of training by considering users of different ages. For this reason, future studies should be done to test the porotype also with the elderly, adolescents, and children to verify the inclusiveness of the experience. Finally, as mentioned earlier in the first discussion (section 2.2.5), some of the guidelines proposed concern specifically the river flood emergency and are difficult to generalize to other situations. In our case, the co-design methodology was applied to a very specific situation, but future research should expand its application to the design of VR simulations of other emergency situations, an area not adequately covered at the moment.

Based on the obtained results, further research will deepen users' behavior during the experience and investigate the link between different variables that impact learning.

3. Study 2 - Shared watching experience and interactive streaming service

3.1 Introduction

Watching movies has always been an activity that has driven spectators to interact, either in presence by commenting on the scene with their fellows or online, by posting reviews and opinions on dedicated blogs. Indeed, watching video programs along with the loved ones is a leisure activity that has proved to foster positive effects on people's health and well-being (Fitzpatrick, 2009; Berg et al., 2001). Additionally, watching TV together can provide several valuable opportunities to strengthen social ties among viewers Cohen and Lancaster (2014).

The shared movie experience is relevant to the emerging field of interactive movies, which allow users to actively participate in the story by making narrative choices that influence the outcome and flow of the plot. This form of cinema, namely interactive cinema, enables a two-way interaction between users and media: the user not only can join the flow of the interactive storytelling, but s/he can also place him/her as a character thanks to the plot-choices provided (D'aloia, 2020). As such, interactive digital narratives (IDN) have the potential to change the way in which we experience cinema and films, and the way in which we engage in storytelling. Indeed, IDN are positively valued by the audience (Roth and Koenitz, 2019).

Nowadays, it is possible to watch these types of media remotely, not only on one's own but also in groups, thanks to specific services offered on some streaming platforms. These platforms provide an extra application that enables remote co-viewing, known as social TV viewing, which adds a social engagement component to what was once an entertainment-only medium (Kim et al., 2021). Indeed, the term social television viewing refers to the simultaneous act of watching the media while communicating with other online users about the same content (Kim et al., 2018). Specifically, with these new streaming services, a group of users can synchronize the media they are watching on a streaming platform (e.g., Netflix) and discuss the show via real-time text chat as if they were all sitting together (Mulla, 2022). A shared visual space can improve communication, as well as task execution and can help to create a common ground. By achieving common ground, however, group members focus more on completing the task rather than coordinating their respective roles (Macaranas et al., 2013). These novel interactive opportunities have been mainly exploited during the Covid-19 lockdowns, when due to health restrictions, social activities, including social gatherings, were forbidden. The pandemic has changed consumer behavior toward content consumption, making social viewing a significant factor in the adoption of streaming platforms (Anthony and Falzon, 2020). In the early stage of the pandemic, three million of UK consumers accessed Video-on-Demand Services, spending approximately one hour and eleven minutes per day on platforms such as Netflix, Amazon Prime Video, and Disney+ (Paun & Olsen, 2022).

The use of online platforms to watch and participate in interactive movies has the potential to create different types of online communication compared to traditional movies. Group viewing of an IDN could change the meaning of participation, enjoyment, and performance of users, who operate together for a common goal bound to the narrative. The dynamics of interaction between users could also change, as they are forced to express opinions and have to choose the flow of the story together. Remote group experience of these films would perhaps need a richer communication modality than the text chat offered so far by streaming platforms. The potential for online communication during digital narratives has been largely unexplored by previous research. Although computer-mediated communication while watching online contents has been studied already (Geerts, 2006; Weisz, 2008), the unique characteristics of interactive movies and their potential to create different types of online communication among viewers make this an area that deserves further investigation.

The present study aimed to investigate group communication dynamics during a social TV experience: watching, synchronously and remotely, an interactive digital narrative. More specifically, we compared three communications modalities, being text chat, audio chat, and audio-video chat. Moreover, the IDN was offered to both participants watching in small groups and on their own, to explore their overall enjoyment of the experience.

The remainder of the paper is organized as follows. The next section will provide a theoretical background, followed by related works. Subsequently, the study will be described, focusing on the experimental method, procedure, and measures used. Following, the analyses will be detailed and the outcomes discussed.

3.2 Background

3.2.1 Interactive digital narratives

Interactive Story-telling (IS) was born in the attempt to make the experience of storytelling more interactive, offering the possibility for users who are watching a media to make meaningful decisions for the story and play an active role in directing it. The origins of Interactive Digital Narratives can be dated back to the 1960s, when video games leaked narrative development and were characterized by extreme simplicity (Kent, 2001). Later on, with advances in technology, users have become more eager to control the events in the game (Klimmt, Hartmann, & Frey, 2007). Developing an interactive video game is a real challenge because the more game-like an experience is, the less impact the narrative will have and, vice-versa, the more the narrative attempts to amaze and entertain the player, the fewer options there would be to manage it (Roth, 2015). Thus, video games began to take different forms and paths, with interactivity being the defining characteristic of the development of this new wave of video games. Specifically, interaction refers to enabling the user to modify the unfolding of the story in real time according to the choices he/she makes (Porteous, Cavazza, & Charles, 2010).

Around the 2010s, the term Interactive Digital Narrative gained a broader meaning, referring to different forms of digital interactivity based on interactive storytelling, such as Interactive Fiction, Drama, video games, or Hypertext Fiction (Koenitz, 2015). One of the reasons why IDN has become so popular, is that it allows the same story to be presented from various perspectives within different cultural and social cause-effect models. Thus, it enables an interactive and participatory process where the individual actively contributes to the narrative.

The interaction with IDNs has been traditionally considered a topic of research for Human-Computer Interaction (HCI), in which the communication between the user and the story is a form of computer-mediated communication (Roth, 2015). Previous studies have investigated the IDN focusing on a single user interacting with the media (Roth and Koenitz, 2019). In the evaluation of the user experience related to IDN two approaches can be identified. The first one focuses on the mere usability of the system, assessing aspects like ease of use, system responsiveness, ability to process the user's input and provide an adequate output, and hardware and software components. The second approach addresses the capability of the system to entertain, thereby including elements related to the narrative, characters, and user's interactions with the system. Roth and Koenitz (2016), tried to integrate these two approaches by creating a revised tool based on the principles of affordances and aesthetic qualities, defining three different macro-areas to evaluate IDN: i) agency, that is the condition of feeling control and implementing changes in a virtual environment; ii) immersion, that is the involvement and active suspension of disbelief; and iii) transformation, which is the effect of the choices made and consequences considering that alternative paths exist, combined with aspects of user experience. They recently tested this instrument with an IDN product created by Netflix, "Black Mirror - Bandersnatch", in a study aimed at identifying the factors that guided single users' enjoyment and experience (Roth and Koenitz, 2019). "Bandersnatch" is an interactive movie that allows the user to make choices on behalf of the main character, leading to different and multiple endings. Roth and Koenitz found that the current Netflix technology and prerecorded material cannot guarantee a complete sense of agency, especially with regard to the ability to influence intelligible changes in the virtual world and feel in control. These results highlight the need for new design strategies to provide audiences with meaningful choices to deliver the best experience (Murray, 2012; Roth and Koenitz, 2019). Finally, the communication dynamics that would emerge during IDN viewing in a group have therefore not yet been investigated and require the co-investigation of the study area related to Computer-mediated communication.

3.2.2 Computer-mediated group decision making

The study of the interactions between people and computers has broadened our understanding of the dynamics of online users, exploring and explaining the concept of the group, particularly in the group decision-making process. In computer-mediated groups (CMGs), where less identifying information is available, users are more likely to be perceived as members of a social group than as individuals on their own (Ruggieri, 2011), increasing their social identity. When social identity becomes salient, depersonalization processes occur (i.e., the user sees him/herself and other group members in terms of group identity). As a result, users tend to perceive the team as a single entity, increasing their commitment, trust, and satisfaction (Tanis & Postmes, 2007). Depersonalized CMGs result in solid identification, group identity salience, and cohesion[5] (Ruggieri, 2011). As regards CMGs involved in leisure activities, one of the main aspects experimented concerns the possibility of making users who connect online perceive the same feeling of the face-to-face experience, in what is called "social television".

It is well known that the social aspect plays a fundamental role in the online viewing experience because they increase the sense of physical presence between people (Shamma, 2008; Harboe, 2008). For example, Cohen and Lancaster (2014), evaluating social media co-viewing, have found that social TV can provide several valuable opportunities to strengthen social ties among viewers who usually watch television alone. Additionally, research on computer-mediated group decisions making has been focused on comparing CMG with face-to-face (FTFs) contexts and has addressed several variables of online interaction (Schlosser, 2009; Baltes et al., 2002). For instance, in group decision-making the order in which alternative courses of action are considered. Indeed, the pieces of information considered first, are likely to orient the entire process that will follow (Kelly 2004). In addition, previous research suggested that when users have to make a decision in group, they are influenced by the majority opinion instead of supporting their own thoughts. In CMG, this effect of social conformity is more evident as the social cues in the group increase (Laporte, 2010). Besides, it was found that the group size was not significant and reported

[5] the pride in belonging to a particular group, interpersonal attraction, and commitment to achieving the goal.

no differences in the quality of decisions made when comparing CMGs with FTF groups (Baltes, 2002). Another aspect that determines high group performance in decision-making concerns trust and relationships among the group members: the higher the trust, the better the performance. CMGs, on the other hand, seem to be more focused on task resolution rather than socialization, possibly because more effort is required to establish stronger bonds (Carte, 2006).

Further analysis showed that the communication channel also affects group efficiency. It has been found that the audio channel was linked to the highest degree of efficiency in CMG. This effect is explained by the fact that in CMC every behavioral cue is processed as a single image. Oppositely, in FTF several behavioral cues are processed as a unique information pattern (e.g., gaze, gestures) (Doherty-Sneddon, et al., 1997).

However, only a few authors have investigated communication modalities linked to the experience of watching online videos in groups. Geerts (2006) compared text and audio chat while watching television programs. The results show that audio chat is generally considered more natural and makes it easier to keep on following than texting. Nevertheless, an inconsistency issue was reported in the overlapping of auditory stimuli (i.e., people's voices, TV audio). Weisz (2007) argued that users' enjoyment would be increased by watching online media synchronously using text chat. The results showed that meeting together for a shared experience in CMC contexts creates a sense of group cohesion: text chat can significantly improve the relationships of the individuals involved in watching the online video. Nevertheless, text chat does not lead to a change in the degree of enjoyment: users who watched online media together had fun with or without chat.

A major obstacle in this research area concerns the overlapping of different attentional stimuli. The visual channel is overloaded through text chat by two stimuli: the text produced and the media playing in the background. Yet, Metcalf and colleagues (2008) reported that people who communicated via pre-set text messages wanted to use richer forms of communication while watching online videos. The results of other research by Weisz (2008) on text and audio chat also highlighted that both communication systems are appreciated, even if there is a preference for audio chat. Besides, the text chat led to more significant distractions not caused by the media itself, rather by the textual conversation. Further research (Macaranas et al., 2013) underlined that also a video chat used to communicate remotely while watching a video online is engaging and fun and improves social ties between participants. Moreover, Baillie (2007) discovered that an audio chat is a useful feature to remotely communicate while watching TV, with a comparable degree of joint TV experience and social presence as face-to-face context. Finally, another study (Schatz, 2008) reported that when a user watches TV alone and communicates remotely using an audio chat with other users, they talk more than when users are actually co-located all in the same room.

However, the effects of these communication systems on group interaction while watching an online interactive video have not been investigated yet.

3.3 The Study

The aim of this study was to investigate how different communication modalities affect group

The aim of this study was to investigate how different communication modalities affect group dynamics and user experience during synchronous IDN viewing. More specifically, we explored the effects of communication modalities on the overall user experience and on the narrative choices and interactions made while watching the IDN in a social TV situation.

More specifically, the study addressed the following research questions: i) How does the communication modality affect the user experience during online synchronous IDN viewing? ii) Which communication modality is most effective in facilitating the decision-making group?

Based on previous findings showing that richer communication systems that do not interfere with the visual channel during media viewing are preferred (Geerts, 2006; Weisz, 2008; Harboe, 2008; Baltes, 2002), we expected that the audio-chat would be the communication modality that would lead to both a better group performance and a greater enjoyability of the overall experience.

The present research was designed in accordance with the Declaration of Helsinki and approved by the ethics committee for psychological research at the University of Padova.

3.3.1 Experimental design

The experiment followed a between-participant design. The only factor manipulated was the type of communication modality, being text chat, audio chat, audio-video chat, or none (absence of a communication system). Participants were randomly assigned to an experimental condition, and they all watched the same episode.

3.3.2 Task

Participants were asked to watch an interactive episode streamed online in small groups. The episode included 19 narrative crossroads, each resulting in a correct or incorrect outcome. The episode belongs to the adventure genre, and the main goal is to keep the main character alive, by selecting the correct alternative course of action in all the narrative crossroads. Only one of the participants had access to the commands (with random assignment within group members).

3.3.3 Episode structure

The episode that was deployed in the present study is called “Lost on Snow Mountain” and it is the 4th of the interactive series “You vs Wild” by Netflix. The episode lasts 18 minutes and opens to 19 different narrative crossroads and 20 different narrative plots. Not all participants would go through all of narrative crossroads, because the order and the number of crossroads change depending on the narrative choices made by the user, each resulting in a correct or incorrect outcome. The narrative crossroad #1 distinguishes two different experiences (Static-Dynamic). The narrative choice made at this point would open up two narrative paths. By choosing the Static storytelling, the user would then be presented 3 further narrative crossroads. Differently, the selection of Dynamic storytelling would bring to 4 more narrative crossroads.

As mentioned above, the episode belongs to the adventure genre, and the main goal is to keep the main character alive, by selecting the correct alternative course of action in all the narrative crossroads. Consistently, the ideal path is characterized by choosing all the alternatives (i.e., Static, Shelter, Fish, Fire) that would make the main character survive, feeding him and giving him all the necessary supplies in the shortest time possible (Figure 3.1).

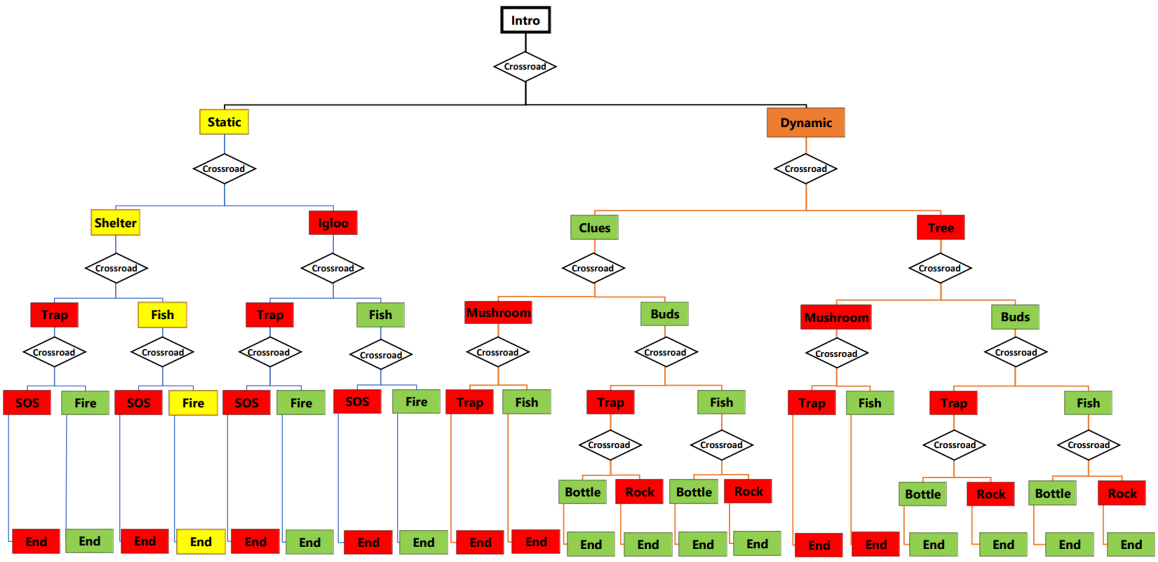

Figure 3.1: Structure of the episode. It is composed of scenes (i.e., rectangles) and crossroads (i.e., rhombuses) with two alternatives. The green alternatives are the correct outcome, and the red ones are the incorrect outcome. The golden alternatives represent the ideal path. The blue lines indicate the static experience and the orange ones indicate the dynamic experience.

3.3.4 Measure

Objective and self-report measures were collected to investigate the effect of the communication modalities on an online group decision making and user experience while watching the interactive episode.

3.3.4.1 Objective data

To assess the user's performance in terms of achieving at making the main character survive, frequencies of correct narrative choices and the selection of the ideal path were collected:

- **Frequency of correct narrative choice:** the frequency of correct narrative choices selected by each participant or group at each narrative crossroad was considered as a proxy of decision making effectiveness. It was grouped by communication modality;
- **Frequency of ideal path selection:** the frequency with which the participant or the group made the all the correct choices was considered as overall efficiency. It was grouped by communication modality;

To understand group involvement, the frequency and contents of group interaction were recorded:

- **Frequency of group interaction:** the frequency of interaction during the experience was assessed as socialization and cooperation;
- **Topics tackled during the interaction:** the topics that participants tackled during the interaction were coded and served as a metric of quality of interaction.

Group aspects related to status equalization (Ruggieri, 2011; Tanis & Postmes, 2007) and social conformity (Kelly, 2004; Laporte, 2010) were assessed by coding the role of the participant (leader vs no-leader) who first proposed a possible option at each narrative crossroads, and the frequency with which such proposal was accepted by other group members:

- **Role prevalence in the first choice proposed:** the frequency with which, at each narrative crossroad, the first possible option was suggested by the leader or other group members.
- **Tendency to social conformity**: the frequency with which the first option proposed by a group member was accepted by the others.

3.3.4.2 Self-report data

We collected information about users' backgrounds and characteristics using a questionnaire in which they were asked to report socio-demographic data (e.g., gender, age, nationality, occupation) and previous experiences with IDN.

In addition, the user experience evaluation while watching an interactive episode was recorded considering the three macro-areas suggested by the literature, being agency, immersion, and transformation (see Section 1 for details). To do this, a post-experience questionnaire was administered; the questionnaire items were adapted from Roth (2015) and created ad-hoc. More specifically, the dimensions investigated are described below.

- **Agency** was addressed by three dimensions: system usability, effectance, and autonomy. The system usability of IDN (5 items) referred to the ease of use, intention of use, perceived time pressure and level of satisfaction with the IDN system. Effectance (4 items) referred to the influence users perceive to have on the narrative and thus also to the awareness of their power to change the narrative within the interaction (Deci et al., 2000). Autonomy (4 items) investigates the sense of independence and freedom, which in an IDN comes from being able to choose from a range of options without feeling compelled to choose a particular path (Roth, Koenitz, 2016).
- **Immersion** included five dimensions: flow, spatial presence, identification with the character, curiosity, and suspense. Flow (2 items) is defined as the optimal experience between the system's challenges and the user's abilities. It is achieved only if the user is found to feel an optimal state of experience intermediate between boredom and anxiety. Indeed, tasks that are too easy can lead to disinterest, while tasks that are too difficult can elicit frustration (Roth, Koenitz, 2016). Spatial presence (2 items) assesses the feeling of being physically present in the story environment. Identification with the main character (2 items) assesses the extent to which the user can take the main character's perspective. Curiosity (5 items) is defined as an internal state that occurs when perceived uncertainty generates a tendency to engage in exploratory behavior to mitigate the uncertainty. Finally, suspense (2 items) is that form of empathic distress perceived by a media user while watching typically dramatic episodes. These aspects are important in assessing emotional involvement toward the proposed narrative.
- **Transformation** was defined by the enjoyment dimension (2 items) that refers to the sense of amusement. It underlies entertainment and it is associated with a feeling of positive experience (Vorderer et al., 2004).

Moreover, aspects related to the psycho-social processes of CMGs were investigated. In particular, participants' perceptions of the occurrence of anti-normative behavior (3 items); social influence (4 items)[6]; intra-group motivation (2 items), group immersion (2 items), were explored. Further, opinions regarding the modality of communication offered were collected in order to check the quality and goodness of computer-mediated interaction and communication. To this end, the pleasantness of the experience, ease of use, originality, and fun related to the

[6] i.e., whether the mere presence of other users is able to elicit a change in the user's ideas

communication modalities were investigated with 5 items. Finally, participants were asked if the interaction with the narrative crossroad had bothered them (1 item).

The CMG aspects were not investigated in the "No communication system" condition, because the user was interacting individually with the episode.

Participants responded to the post-experience using a 5-point Likert scale (1= totally disagree; 5= totally agree).

3.3.5 Experimental procedure

The experiment was conducted online and participation was voluntary. Only individuals with an active Nextflix account were eligible to participate in the experiment. Moreover, they had not to have watched the episode yet, and have access to a desktop PC or a laptop to play the experience. Once participants had agreed to participate, they received an e-mail with the date and time of the online experiment with a link to connect to the Zoom meeting software (version 5.10.7)[7].

Participants were randomly assigned to one of the four experimental conditions. Those assigned to one of the three group conditions were asked to recruit two friends with a Netflix membership, with whom they would simultaneously watch the episode online using the Teleparty service. Therefore, each group consisted of three participants who already knew each other and could communicate through the Zoom platform. The control of the commands (i.e., move the mouse and click on the preferred alternative) was randomly assigned to one participant in each group.

On the day of the experimental session, participants first logged into Zoom to start a group video chat; they were asked to read and accept informed consent and filled out a demographic questionnaire. Before starting the interactive episode on Netflix, the experimenter provided instructions on how to use the interactive interface to make their narrative choice at the crossroads, the goal of the episode, and the communication rules. Participants were informed that they could ask for clarification at any time during the experiment. In group conditions, each group could communicate using the assigned system (i.e., text chat or audio or video chat) offered by Zoom.

The experimenter followed the experimental session via the Netflix TeleParty and Zoom, to track the choices made and the interactions. Moreover, the session was recorded using the OBS Studio software (version 27.2.4)[8]. Once the episode ended, participants were asked to turn off the camera and microphone and individually fill out the post-experience questionnaire administered using the online survey system Google Forms[9]. The experiment lasted about 35 minutes.

[7] https://zoom.us/

[8] https://obsproject.com/

[9] https://www.google.it/intl/it/forms/about/

3.3.6 Participants

A total of 120 participants took part in the study, equally distributed for each experimental condition. An a priori power analysis conducted on Gpower (Erdfelder et al., 1996) indicated that a total sample of 120 people was needed to detect a medium effect size (d = .5) with 95% power. 30 participants (13F) aged 21 to 29 (M = 24.73, SD = 1.85) were assigned to the single player (No communication system) condition while 30 participants were assigned to each group conditions (i.e., 10 groups of 3 individuals each) respectively: text chat with 30 participants (F =20; Mage = 25.20; SDage = 4.04; range 20-33); audio chat included 30 individuals (F = 14; Mage = 23.30; SDage = 2.33; range 18-28); and finally, the video-audio chat comprising 30 persons (F = 22; Mage = 22.93; SDage = 1.89; range 19-26). All participants were native Italian speakers, Netflix account owners, had a good internet connection, and basic computer skills.

3.4 Analysis and results

Video recordings were analyzed offline using dedicated software, namely the Boris Software (Friard & Gamba, 2016) following a specific coding scheme (i.e., top-down analysis), meant to assess performance, group dynamics, status equalization and social conformity. In the pre-processing phase, the comments related to technical problems (i.e., Wi-Fi connection) that arose during experimental sessions were removed from the total number of comments produced.[10]

For what concerns questionnaires (i.e., demographic data, user experience evaluation), the scores of the items with negative phrasing were reversed. The statistical analysis of the collected data was conducted using the R software (R Core Team, 2016).

3.4.1 Distribution of the narrative choice

We used a binomial test to analyze the proportion of correct narrative choices across the different experimental conditions to assess performance. The single-user condition was used as a control condition. Overall, a strong preference emerged for the correct alternative (62% correct response and 38% incorrect response), indicating that the selection of the correct alternative was not random. Then, we compared each of the experimental group conditions with the control one, using the probability of success in the control condition as the expected value for the null hypothesis. The null hypothesis was rejected only for the audio chat condition ($p < .001$, probability of success = .86), indicating the group in which the interaction was mediated by audio

[10] These problems did not invalidate any of the experimental sessions because the comments were sporadic.

was very efficient. No significant preference for the correct alternative was observed in the text chat ($p = .118$, probability of success = .69) or video-audio chat ($p = .057$, probability of success = .53) conditions (Table 3.1).

Distribution of the correct choices compared to the no support condition					
Group condition	**A - Correct**	**B - False**	**p-value**	**Estimated P "success"**	**Null hypothesis H0**
Text Chat	87 (69%)	39 (31%)	.11	.69	.62
Audio Chat **	108 (86%)	18 (14%)	< .001	.86	.62
Video-Audio Chat	69 (53%)	60 (47%)	.05	.53	.62

*Table 3.1: Value of binomial tests, p(B.H.), frequency and percentage of correct and incorrect narrative choices, p-value, Estimated P "success" and Null hypothesis (**p < .001).*

3.4.2 Preference for the ideal path

We investigated if there were differences in the choice of the ideal path under different conditions. A comparison was made between the number of participants who took the ideal path and the number of participants who took other paths for each condition. The ideal path was taken by 9 participants in the text chat condition (3 groups), 18 in the audio chat condition (6 groups), 3 in the video-audio chat condition (1 group), and 5 in the no support condition. The participants' choices of the group conditions were compared with the single-user condition, used as a control, to estimate if the communication system affected the choices.

A Fisher test revealed a difference in the frequency of choosing the ideal path for the audio chat (18/30; $p = .001$) compared to the control condition. No difference emerged between the single-player condition and the text (9/30; $p = .36$) and video-audio chat (3/30; $p = .71$) conditions. Finally, another series of Fisher tests was conducted to investigate whether differences between the different group conditions. The aim was to check whether the audio chat was more likely to lead to the ideal path than the text and video-audio chat.

Audio chat vs. text chat. A difference was shown ($p = .03$, $OR = 3.42$). Participants in the audio chat were more efficient in reaching the goal (Ideal Path = 18) compared to the ones in the text chat (Ideal Path = 9).

Audio chat vs. video-audio chat. The two conditions differed ($p < .001$, $OR = 12.84$). A lower number of Ideal Paths ($n = 3$) were shown in the video-audio chat compared to a higher number in the audio chat ($n = 18$). Participants seemed to be more efficient in reaching the goal than the video-audio chat ones (Figure 3.2).

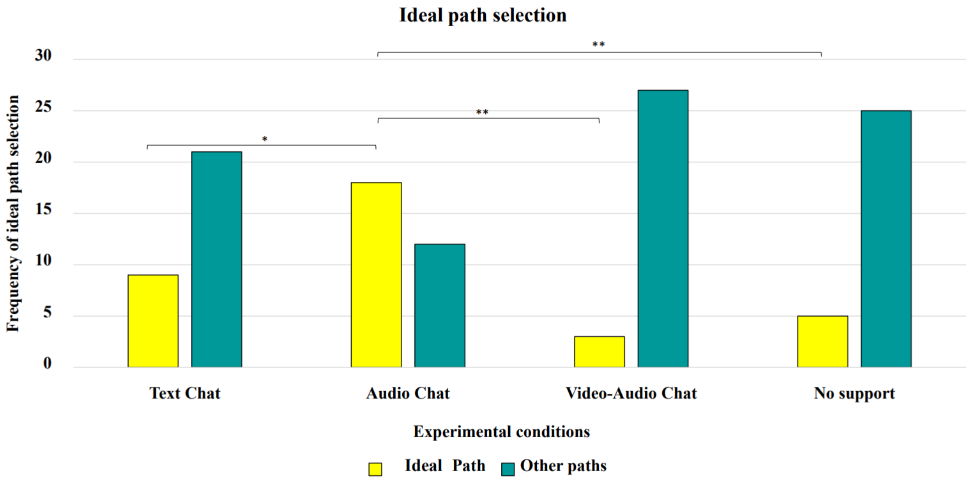

*Figure 3.2: Distribution of the ideal path selection for each condition (*p < .05, **p < .001).*

3.4.3 Distribution of comments produced during the episode

A linear regression was conducted to evaluate potential differences among conditions (text chat, audio chat, and video-audio chat) in terms of comments produced by each group while watching the episode. Audio chat and video-audio chat were compared to text, as it represents the only system currently offered in streaming services. The mean production of comments was higher ($t = 2.25$, $p = .03$) in the audio chat ($M = 108.60$, $DS = 62.51$, $Mdn = 91.00$) compared to the text chat ($M= 65.10$, $DS = 42.73$, $Mdn = 58.50$). Besides, another difference was highlighted ($t = 2.08$, $p = .04$) in favor of the video-audio chat ($M = 114.80$, $SD = 51.58$, $Mdn = 104.50$) compared to the text chat (Figure 3.3).

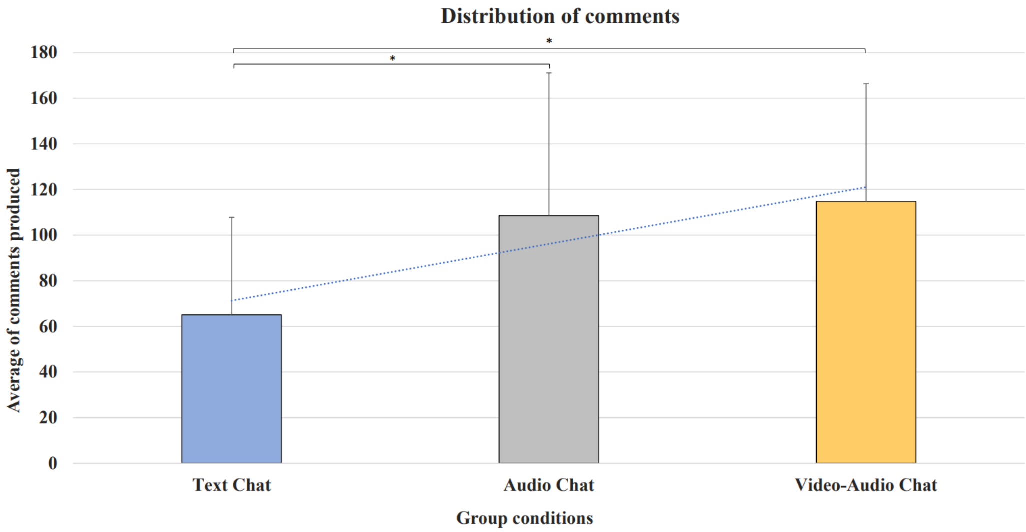

*Figure 3.3: Distribution of comments in Text Chat, Audio Chat, and Video-Audio Chat conditions (*p < .05).*

3.4.4 Comments production between narrative crossroads and scene phases

The analysis concerned the proportions (%) of comments linked to narrative crossroads vs scenes phases. A series of beta regressions were carried out. Based on the lowest AIC the model that best fitted the distribution of the data was identified.

The regression model used (*phi* = 3.21, s.e. = .52, *z* = 6.11, p < .001) showed a difference (*z* = -4.25, p < .001) between the narrative crossroads and the scenes comments produced during the episode respectively in text chat (narrative crossroads: 56%, scenes: 44%), audio chat (narrative crossroads: 62%, scenes: 38%) and video-audio chat narrative crossroad (narrative crossroads: 66%, scene: 34%) conditions, highlighting more commentary production during the narrative crossroads. (Figure 3.4).

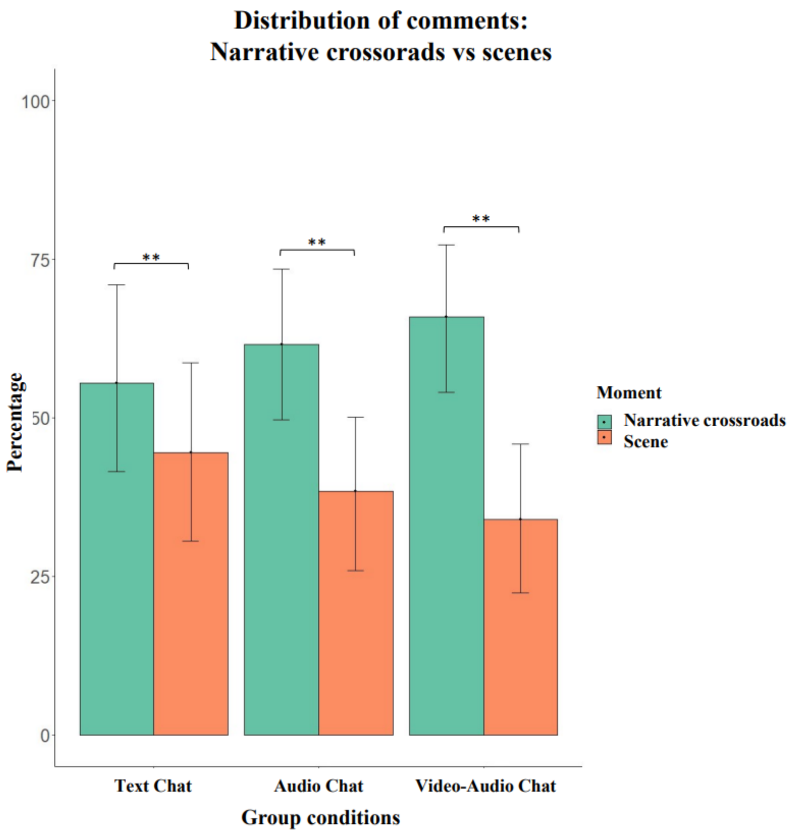

*Figure 3.4: Percentage of comments produced while watching the episode for each group condition in the narrative choice and during the scene (** p < .01).*

3.4.5 Contents of comments at the narrative crossroad

A thematic analysis approach was used to analyze the contents of the comments made at the narrative crossroads. Each comment could fall into one of the following categories: proposal (identifies a suggestion put forward for consideration by other group members); agreement (expression of approval by one member related to an action proposal); disagreement (lack of consensus regarding an action proposal by at least one group member); indecision (in how to make a proposal or choose between different alternatives); doubt (general expression of uncertainty); encouragement (invitation addressed to one group member to say/choose something); argumentation (provision of motivations in support of one's position); off topic (a comment unrelated to narrative choices); objective (expression of ideas related to task); laugh (sound of laugh in the audio chat and video-audio chat condition or smile emoticon in text chat condition).

A series of Beta regressions were carried out. The "proposal" comments in the text chat were taken as the baseline, because proposals initiate and guide the decision-making process and enable to assess the quality of the interaction and the psychosocial phenomena occurring within groups (Kelly, 2012). The analyses were carried out taking into consideration the proportion (%) of comments made during the narrative crossroads. The text chat was chosen as a reference as it represents the current alternative in streaming services (Netflix, Prime, etc.). The regression model (*phi* = 7.33, s.e. = .680, z = 10.79, p < .001) showed differences both in the comparison between the comments in the text and in the audio chat (z = 2.66, p = .02) and in the text and video-audio chat (z = 2.79, p = .01). In the text chat condition participants mainly made more proposals while laughing and arguing less.

In the audio and video-audio chat conditions, participants, in addition to proposing the alternative to be selected, tend to argue each proposal more, express their degree of agreement and disagreement more frequently, and laugh more (Table 3.2).

Content distribution of comment			
Comment content	Text chat	Audio chat	Video-audio chat
Proposal	50.32%	18.00%	16.59%
Agreement	14.16%	20.87%	18.79%
Disagreement	1.49%	.76%	1.38%
Indecision	3.89%	2.71%	2.13%
Doubts	5.05%	8.23%	5.34%
Incitement	8.97%	7.39%	8.27%
Argumentation	14.30%	29.95%	29.11%
Non-inherent	.49%	5.65%	6.45%
Focus on the goal	.76%	.93%	.73%
Laugh	1.02%	5.72%	11.37%

Table 3.2: Percentage of the different contents of comments produced during the narrative crossroads.

The model also showed that, in all the group conditions, there was a difference between the production of proposals made with: degree of agreement (z = -3.15, p = .001); and disagreement with the proposal (z = -10.16, p < .001); indecision regarding the proposal (z = -8.48, p < .001); doubt about the proposal (z = -7.36, p < .001); incitement to follow the proposal (z = -6.01, p < .001); comments not related to the crossroad (z = -8.30, p < .001); comments regarding the goal to be achieved (z = -10.66, p < .001); production of laughs at the time of choice (z = -7.809, p < .001).

3.4.6 User experience evaluation questionnaire

User feedback: comparison between the experimental conditions. The distribution of the data was not normal so nonparametric analyses were performed. A series of Kruskal-Wallis tests were conducted to assess potential differences in the user experience. For the no support condition, the dimensions related to the group experience were not investigated. In case of significance, a series of post-hoc comparisons were conducted with the Mann-Whitey test and p values adjusted with the B.H. method. Below are reported results that showed a statistically significant difference in Kruskal-Wallis tests, while all the results obtained can be consulted in Table 3.3.
The dimensions of ease of use ($\chi 2$ = 11.99, df = 3, p = .01), and UX communication system ($\chi 2$ = 8.124, df = 2, p = .01) highlighted differences across conditions. A series of Mann-Whitney tests were carried out. Considering the ease of use, a difference emerged between the no support and text chat (p = .01). In the single-player condition the choice system was perceived as easier to use (Mdn = 5.00, M = 4.95, SD = .20) compared to the text chat (Mdn = 5.00, M = 4.66, SD = .49).
Regarding the UX communication system, the comparison revealed differences respectively between the text and audio chat (p = .02) and between the text and the video-audio chat (p = .02). The scores of those who used text chat (Mdn = 3.00, M = 3.36, SD = 1.03) were lower compared to the scores of participants who communicated via audio (Mdn = 4.00, M = 3.93, SD = 1.01) and video-audio chats (Mdn = 4.00, M = 4.03, SD = .93); (Table 3.4).

Post-experience questionnaire														
Dimension	**No support**			**Text chat**			**Audio chat**			**Video-Audio chat**			**Kruskal-Wallis tests**	
	Mdn	**M**	**SD**	**Mdn**	**M**	**SD**	**Mdn**	**M**	**SD**	**Mdn**	**M**	**SD**	**X^2**	**p**
Ease of use*	5.00	4.95	.20	5.00	4.66	.49	5.00	4.83	.46	5.00	4.88	.31	11.99	***.01**
Intention of use	4.00	3.46	1.33	4.00	3.73	.90	4.00	3.76	1.22	4.00	3.70	1.05	.87	.83
System satisfaction	4.00	4.23	.68	4.50	4.36	.76	4.00	4.30	.53	4.00	4.26	.78	1.06	.78
Time pressure felt	4.50	4.03	1.25	4.00	3.66	1.32	4.00	3.50	1.16	4.00	3.66	1.21	4.34	.22
Autonomy	2.50	2.93	1.14	3.00	3.05	1.02	3.25	3.15	.91	3.25	3.06	1.08	.66	.88
Effectance	3.50	3.35	.98	4.00	3.90	.80	3.50	3.61	.55	4.00	3.66	.84	5.94	.11
Flow	4.00	4.01	.77	4.00	4.05	.60	4.00	3.85	.64	4.00	4.18	.56	4.35	.22
Identification with the main character	2.75	2.75	1.03	2.50	2.73	.82	2.75	2.63	1.09	2.50	2.60	1.00	.50	.92

Curiosity	5.00	4.43	.72	5.00	4.70	.53	5.00	4.46	.62	5.00	4.53	.50	3.51	.32
Suspense	3.75	3.33	1.15	3.75	3.58	.97	2.50	2.90	1.26	3.00	2.96	1.26	6.47	.09
Spatial Presence	3.00	3.06	.97	3.00	3.13	.99	3.00	2.85	.99	3.00	2.80	1.10	1.56	.66
Enjoyment	4.25	4.13	.89	4.50	4.40	.59	4.25	4.23	.59	4.00	4.35	.54	1.53	.67
Anti-normative behavior	NA	NA	NA	2.00	2.00	.94	2.00	1.86	.73	2.00	1.76	.85	1.10	.57
Social influence	NA	NA	NA	2.75	2.61	1.25	3.00	2.83	1.20	2.75	2.71	1.24	.73	.69
Intra-group motivation	NA	NA	NA	4.00	3.95	.87	4.00	4.10	.72	4.50	4.06	.85	.50	.77
Group Immersion	NA	NA	NA	4.00	3.60	.88	3.50	3.13.	1.02	4.00	3.55	1.21	3.95	.14
UX communication system*	NA	NA	NA	3.00	3.36	1.03	4.00	3.39	1.01	4.00	4.03	.93	8.12	.01
Interaction during scenes	NA	NA	NA	5.00	4.82	.49	5.00	4.86	.45	5.00	4.87	.34	.23	.89

*Table 3.3: Post experience questionnaire: comparison related user evaluation in the different experimental conditions. Value about descriptive analysis (median, mean, and standard deviation) for each experimental condition and Kruskal-Wallis Test score (*p (B.H.) < .05). NA = Not Applicable as dimension related to the group experience.*

Preferences for group or single-player experience. We performed a series of Wilcoxon tests (B.H. correction) to evaluate if users preferred the experience of watching the IDN more as single-player or in a group. Participants in the single-player condition did not report a preference towards a specific experience (V = 132.50, p = .74; Mdn = 3.00, M = 2.80, SD = 1.34). Differently, participants in the group conditions reported a preference: participants in the text chat (V = 270.00, p = .004; Mdn = 2.00, M = 2.16, SD = 1.53), video-audio chat (V = 395, p < .001; Mdn = 2.00, M = 1.86, SD = 1.00), and audio chat (V = 395.00, p < .001; Mdn = 1.00, M = 1.60, SD = .83) conditions choose group viewing episode experience.

3.4.7 Group psychosocial effects

Status equalization. An analysis was carried out to investigate a potential distinction between social roles within the groups interacting online or the occurrence of the status-equalization phenomenon (Dubrovsky et al., 1991). The frequencies with which participants first expressed an option to follow at the narrative crossroads were collected. We analyzed whether the person who first proposed the choice was the leader (the one who was in control of the commands, defined as P1) or the other group members (P2 and P3). No differences emerged. The speaking turn among group members was equally balanced in the majority of the conditions (respectively: p = .87 in text and chat conditions) with the exception of the video-audio condition (p = .03, probability of success = 33%), where a tendency to propose the alternative first was actually in favor of P2 and P3 (67%) and not of P1 (33%); (Figure 3.5).

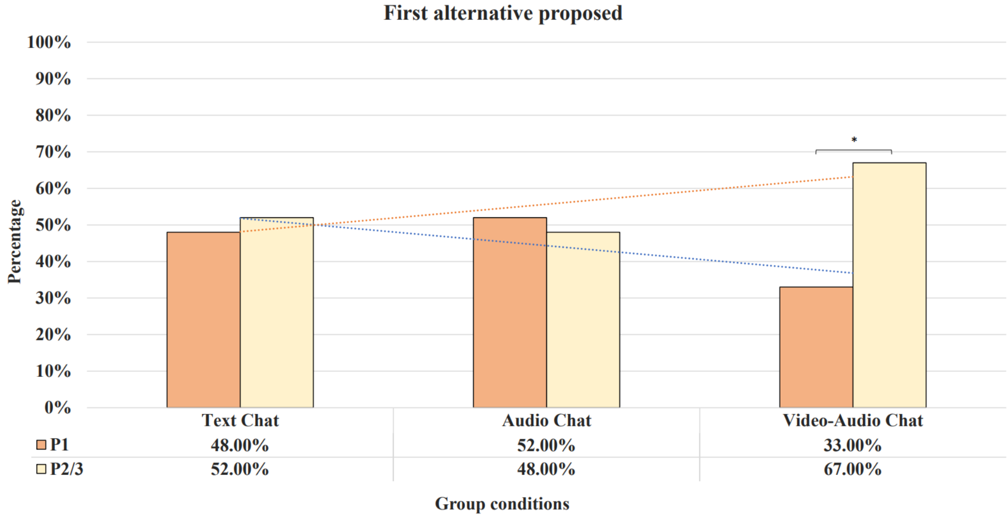

*Figure 3.5: Percentage of user who proposes the first alternative in group experimental conditions (*p < .05).*

Social conformity. The last analysis aims to investigate whether, within each group at the time of choice, the user's discussion has confirmed/changed the first choice proposed by one of the users. Frequencies of confirmation of the proposed first choice were analyzed using a binomial test. In the text ($p < .001$, probability of success = 83%), audio ($p = .002$, probability of success = 73%), and video-audio chats ($p < .001$, probability of success = 79%), there is a strong tendency not to change ideas and confirm the first proposed choice. Those results indicate a kind of social conformity linked to the first moment of the choice (Figure 3.6).

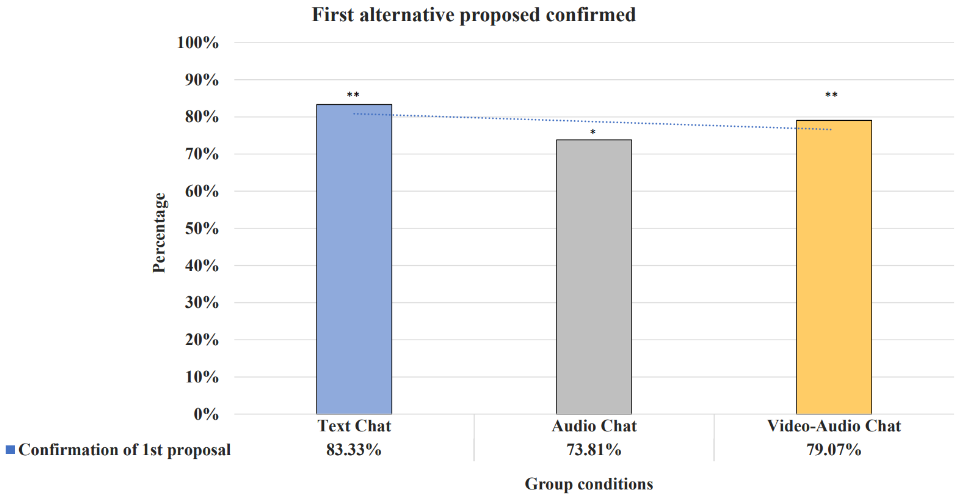

*Figure 3.6: The percentage of confirmation or change of the proposed first choice (*p < .05; **p < .001).*

3.5 Discussion

In this study, we investigated the effect of different modalities of group computer-mediated communication in a social TV situation.

Previous studies addressed communication group dynamics while participants were watching not-interactive TV shows (Geerts, 2006; Weisz, 2007; Schatz, 2008). Differently, here we investigated the effect of an interactive digital narrative on communication group dynamics. To this end, we compared different communication systems (text chat, audio chat, audio-video chats, and no communication system) during a synchronous viewing of an interactive adventure episode. More specifically, we studied the impact of the communication modality on the likelihood of keeping the main character alive, which was the goal of the episode, and how the communication modality affected the decision making process within the group.

Our data indicate that audio only is the most effective communication modality to achieve the narrative goal in an interactive media. This finding confirms our hypothesis and extends previous research that was limited to a traditional group context (Löber et. al, 2006). In particular, not only groups that interacted via audio chat had 86% likelihood of selecting the correct alternatives, but they also had a highest likelihood (60%) of selecting the ideal path.

Notably, our sample was composed only by members tied by genuine friendship bonds, thereby indicating that their experimental choices were independent from the level of trust in the group, as suggested by previous research Carte (2006).

In all conditions, most comments were produced when participants reached a narrative crossroad and a decision had to be made. This result highlights that the salient moments of narrative trigger the discussion within the group. In addition, computer-mediated audio and video-audio chat communication modalities promoted more group interaction. Indeed, users interacting with these two communication modalities tended to produce almost twice as many comments as text chat. Moreover, after an alternative option was proposed, participants tended to argue and express their level of agreement to a higher extent. Nonetheless, comparing those two communication systems, in the audio only condition the social status of group's members was more equally distributed. More specifically, in the audio condition, any member of the group equally proposed first an option to follow, at the narrative crossroads. In the video-audio condition, instead, the tendency to propose the alternative to follow first, was actually in favor of group members who weren't in control of the commands.

The equalization effect leveled out any discrepancies in the users' status without them feeling angry or annoyed (as reported by the results of the dimension "anti-normative behavior"). It was additionally proven that, across conditions, at the narrative crossroads, there was an effect of social conformity: users tended to confirm the first alternatives proposed, in line with the theory postulated by Kelly (2004). In the Attentional Focus Model (AFM), Kelly argued that the group uses initial preferences expressed by a member as an attentional filter to evaluate the subsequent alternatives and determine decision-making processes.

Therefore, the communication modality does not affect the user experience with the IDN. Overall, participants enjoyed the interactive episode, regardless of the communication systems they experienced, in line with previous research (Roth, 2019). It also emerged that users who participated in groups were satisfied with their experience and would not have preferred a single-player one. This preference - which emerged especially among users who used audio and video-audio chat to communicate - confirmed also that sharing leisure activities have positive effects on people (Fitzpatrick, 2009; Berg et al., 2001; Cohen and Lancaster, 2014). However, while this multimedia is on one hand engaging, on the other hand this genre of story makes it hard to identify with the protagonist and does not allow users to feel physically immersed in the story. Finally, participants did not report high levels of perceived autonomy concerning the choice system, however high levels of autonomy do not necessarily imply a pleasant user experience (Roth & Koenitz, 2016; Deci and Ryan, 2000).

Nevertheless, users who communicated via text chat did not evaluate this communication modality positively, as they considered it unoriginal and unpleasant. Differently, positive scores emerged for audio and audio-video modalities, indicating a preference of users towards systems that include audio channels.

3.6 Guidelines for the design of interactive streaming services

Communication system guideline

- provide users with the possibility of communicating throughout the duration of the interactive film (and not only at the moment of the narrative crossroad)
- implement communication modalities that provide for the possibility of interacting via audio

Guidelines for the narrative structure

- make users autonomous by inserting different choice alternatives to give users a sense of autonomy

Storytelling guidelines

- enrich the narrative with elements that help to increase the identification with the protagonist.

3.7 Conclusion

While watching synchronous IDN in online and synchronous group, audio and video-audio chats support a more frequent and better interaction between users, especially at the narrative crossroads. In addition, these CMC were perceived more positively than text chat. Besides, it should be noted that the audience who communicated using the audio channel performed better than the other groups and single-player, completing the episode by selecting the ideal path in most cases. Moreover, users who have experienced IDN as a group reported the intention to share more experiences like that in the future. Indeed, computer-mediated interaction systems studied seem to be related to positive phenomenons during IDN experience, such as status equalization effects, absence of flaming, high degree of enjoyment, and social conformity (Kelly 2004; Laporte, 2010). The findings of this study not only broaden the knowledge related to computer-mediated interaction during IDN watching, but also provide design suggestions for current streaming services to improve the performance and enjoyment of an IDN group experience. It is known that social TV provides an opportunity to build viewer engagement with programs, movies, and their advertisers (Nagy & Midha, 2014; Pynta et al., 2014). Our results highlight the importance of offering the possibility to communicate with other users by exploiting an auditory communication system rather than a visual one. This service can be easily implemented on nowadays streaming platforms and future applications thanks to the microphone integrated into the users' devices or the recent remote controls for TV. It is also essential to create contents that lead to a higher capacity of identification with the main character and design IDN taking into account a more comprehensive system of narrative choices and crossroads to increase the sense of autonomy of the users (Roth, 2015).

This study only focused on one TV show genre. Future research could explore computer-mediated group experiences by considering IDN with other characteristics (e.g., different number of alternatives, genre, duration). Moreover, we involved a group of young adults. Future investigations could focus on different age groups.

4. Study 3 - Retail and 3D web and AR

4.1 Introduction

Nowadays, the shopping experience is not limited to the physical store within fixed opening hours anymore, rather it can be done at anytime from anywhere. Online retail shows continually growing numbers. From 2017 to 2021, the profits generated by online retail worldwide increased from 2.2 trillion to approximately 3.8 trillion of US dollars. Moreover, the percentage of online purchases has increased from 12.2% to 22%, indicating a significant change in consumer purchasing habits. Likewise, in Europe from 2017 to 2021, the sales revenue from e-commerce increased from 370 billion to 670 billion euro, and the percentage of online purchases from 9% to 15% (Statista, 2021). The restrictions imposed by the COVID-19 pandemics have additionally emphasized these trends (Dumanska et al., 2021; Jílková & Králová, 2021).

However, purchasing online lacks many sensorial stimuli (e.g., touching the fabric of a shirt) that make the shopping experience so rewarding. Technologies such as 3D web and Augmented Reality (AR) have the potential to overcome these limitations by compensating the lack of sensory stimulation with high-fidelity simulations (Hilken et al., 2022).

From this perspective, it becomes important to realize e-commerce platforms that leverage the strengths of 3D web and AR to enhance the experience, not only of the consumer but also of the seller. To this end, it is crucial to understand end users' needs and preferences.

The first data collection that we present in this study reported both the factors that either facilitate or hinder shop owners to launch and manage their own e-shop as well as the perspectives of trade fair operators to integrate AR into trade shows.

Many companies are progressing acknowledging the necessity of investing in AR to maintain their leadership position and to empower their brand (Leena et al., 2017). For example, Coca-Cola, McDonald's, and General Electric integrated AR in their marketing strategies for creating interactive advertisements and by providing immersive, engaging, and innovative experiences (Scholz et al., 2016). Moreover, Amazon has recently introduced AR into its platform through its 'View in your room' feature (Romano et al., 2021), which allows users to project digital versions of products into their surroundings such as furniture and appliances, to assess their appearance in terms of size, shape, and color. Similarly, IKEA has integrated AR into its mobile app "IKEA Place" to enable customers to virtually place furniture in their homes, limiting their uncertainty about how a new product will fit into their existing arrangements and designed spaces (Ozturkan, 2021; Sholz et al., 2016).

AR represents a starting point for further advances in retailing, offering both sellers and consumers new opportunities for interaction (Farshid et al., 2018) and exploring hybrid words. Head-mounted displays (HMD) allow the user not only to view the objects but also to manipulate them, by interacting directly with the product (e.g., by rotating it, enlarging it, shrinking it, or exploding it into its various components). Studies have shown that interaction with the product increases the perceived pleasantness of the product itself, provides more information about the product, and consequently positively influences purchasing behaviour (Kang et al., 2020; Xu & Sundar, 2015; Wodehouse & Abba, 2014; Ozok & Komlodi, 2009).
In particular, the use of AR helmets, such as HoloLens, has shown great potential in providing consumers with an enhanced shopping experience. AR helmets allow users to see virtual objects as if they were part of the real world, providing a more realistic and engaging shopping experience. Users can interact with virtual objects through gesture-based interfaces, allowing them to manipulate and explore products naturally and intuitively without controllers (Kangas et al., 2022). However, despite AR gaining increasing success, to date, few studies in the literature have investigated the usability and the user's experience when interacting with virtual 3D objects, as well as the possibility of providing different modalities of interaction based on the peculiarities of the object being manipulated. In this regard, the study conducted by Pham et al. (2018) on gestural usability in HoloLens is a good point of reference since it suggests that the size of the virtual objects, their distance from the user, and the scene in which they are displayed influences and affects the gesture elicited. Indeed, objects may have different characteristics, such as color, brightness, shape, and size. For example, the user may need to shrink a product that in its real size is very large (e.g., the size of a car) or enlarges very small objects (e.g., the size of a smartwatch), or, again, display the object in its real size on a 1:1 scale.
In this framework, the second data collection of this study aimed to investigate the user experience of gesture-based interaction techniques with 3D virtual objects that differ in size. To do this, a Microsoft HoloLens 2 AR helmet was used in a shopping experience. During the purchase, participants manipulated virtual products of small sizes and large sizes through hand gestures by two different modalities. One manipulation modality offered the possibility to explore products by changing their dimensions on a manual free scale (i.e., free scaling), and the other one to explore products in their real physical dimensions on a 1:1 scale (i.e., automatic scaling).
The two data collection are described in the following sections. Particularly, in the next session (section 4.2) is described the explorative data collection on 3D web and AR for displaying and

selling goods[11], while the data collection related to customer experience with 3D virtual objects in AR shopping is reported in section 4.3.

4.2 Rethinking the shopping experience: explorative study on 3D web and AR for displaying and selling goods

4.2.1 Background

Previous studies have investigated the factors that encourage or hinder the adoption and use of e-commerce for different business sectors. Wymer and Regan (2005) addressed the factors favoring or inhibiting the adoption of e-commerce by small and medium-sized enterprises (SMEs). Specifically, they identified the following as the most significant antecedents of adoption: need innovativeness, competitive pressure, value, government, reliability, e-commerce technology, models, prior experience, executive experience, cost, priority, security, capital, market, partners/vendors. Hong and Zhu (2006) reframed these factors in a study regarding the adoption of e-commerce systems by companies. They found that for companies the main predictors of adoption were technology integration, web functionalities and web spending, while firm size, partner usage, electronic data interchange usage, and perceived obstacles were found to negatively influence the migration to e-commerce. Concerning B2B, Sila (2013) found that scalability was the biggest contributor to B2B e-commerce usage.

Nowadays, 3D web, VR, and AR technologies provide retailers with new opportunities to enhance their online sales channels and are likely to represent the next step in the development of online shopping. Several studies investigated the effects of these new technologies on consumers' online purchasing habits. For instance, 3D virtual worlds have been shown to lead to a greater sense of telepresence and social presence, and a greater perception of authenticity, factors that increase users' online trust and purchase intent (Peng & Ke, 2015). Moreover, virtual and immersive worlds are perceived as more attractive and appealing (Pantano & Servidio, 2012), and evaluation of user experience within these virtual environments revealed increased engagement and enjoyment. These factors are positively correlated with user satisfaction, leading to increased intention to purchase the product (Papagiannidis et al., 2013). Thanks to the evolution of 3D web

[11] This work is published in Orso, V., Portello, G., Pierobon, L., Bettelli, A., Monaro, M., & Gamberini, L. (2022). RETHINKING THE SHOPPING EXPERIENCE: A QUALITATIVE EXPLORATIVE STUDY ON 3D WEB AND AR FOR DISPLAYING AND SELLING GOODS. In *16th International Conference on Interfaces and Human Computer Interaction, IHCI 2022, and 15th International Conference on Game and Entertainment Technologies 2022, GET 2022-Held at the 16th Multi Conference on Computer Science and Information Systems, MCCSIS 2022 (pp. 142-148). IADIS Press.*

and AR technologies, virtual and immersive stores can therefore represent a starting point for further advances in online retailing, offering both sellers and consumers new opportunities for interaction (Farshid et al., 2018).

4.2.2 Materials and methods

4.2.2.1 Definition of user categories and requirements

Two groups of stakeholders were involved, namely: retailers and trade fair operators (exhibitors and organizers). Eligible participants in the retailer group had to be shop owners or employed as shop assistants. Similarly, for the trade fair group, candidates had to have professional experience as exhibitors or as organizers of trade fairs.

Participants were reached through the researchers' network of acquaintances and search on professional social networks and on the websites of the Chambers of Commerce. The initial approach was made by telephone and consisted of explaining the purpose and modalities of the meeting, and finally inviting them to the interview. Afterwards, participants received an email with the details of the activity, data collection and storage. They also received an informed consent form, which the participant was invited to sign and return. Finally, an appointment for the interview was arranged.

4.2.2.2 Participants

A total of 26 people operating in the retail sector. Only 12 of them (two women) agreed to participate in the data collection, with a response rate of about 46%. The final sample had an average age of 31.3 years (ranging from 24 to 43 years). They dealt with diverse products, including, hi-fi, food, flowers, ironmongery, body care, books, furniture, jewelry, clothing, household appliances and telephony.

For the recruitment of trade operators, a total of 27 people was contacted with a 22% response rate. The final sample consisted of 6 fair operators (one female), whose average age was 44 years (with a range between 27 and 65 years). Participants dealt with professional cooking machinery, professional sous-vis machinery, the production of automation systems, and the production of pieces of furniture (chairs). One participant was responsible for the organization of trade fairs in the field of tourism promotion (organizing about two fairs a year) and, finally, one dealt with the organization of trade fairs events in the jewelry sector (organizing about 50 fairs a year). In addition, a stand designer and fitter were interviewed.

4.2.2.3 Semi-structured interviews

To deeply explore participants' experience, semi-structured interviews were conducted (McIntosh & Morse, 2015). Two different scripts were devised, one for each group of participants. In addition, two short videos (prompts) were prepared for illustrating the technologies of interest, namely 3D web and AR. By doing this, it was possible to introduce and explain the technologies to all participants in the same way, thereby ensuring that they all had the same level of knowledge. The interview script addressed to retailers the topics to be explored were their professional experience (i.e., type of business, goods handled), their experience with online selling, the presentation style of the products in their online shop (i.e., what kind of information was used when a product is placed in the online shop) and the frequency with which the online shop was updated (if any). Finally, the researcher showed the video and explored with the participant the interest and willingness to invest resources in the 3D web.

Similarly, for trade fair operators, the script covered the following topics: professional experience (i.e., years of activity, number of trade fairs attended), the way in which products were typically displayed on the stand and the way in which the stand was organized, the decision-making process for the selection of products to be exhibited during events, the meaning of the fair and the measures taken to maintain the continuity of fairs during lockdowns, the general differences before and after the outbreak of COVID-19 and whether there was any intention to continue with digital fairs after lockdown. Finally, after watching the clip, the participants were asked about their impressions of AR applied to the display of products in trade fairs.

Additionally, an ad hoc questionnaire was devised to investigate the participants' technological experience and was administered via the Google Modules platform. The questionnaire consisted of 14 items in total (to be answered on a 5-point Likert scale) and investigated the following dimensions: personal innovativeness, reflecting individual readiness and willingness to adopt new information technologies (Elie-Dit-Cosaque, Pallud & Kalika, 2012); technology anxiety, which refers to the feeling of concern related to the use of a technology (Spagnolli et al., 2014); perceived relative advantages, the extent to which a technology is perceived to bring benefits (Van Ittersum & Feinberg, 2010); perceived compatibility, which is the extent to which a technology is perceived as consistent and compatible with experiences and needs (Van Ittersum & Feinberg, 2010); perceived complexity, which is the extent to which a technology is perceived as complex to understand and use (Van Ittersum & Feinberg, 2010); perceived trialability, the degree to which users are willing to try a new technology (Van Ittersum & Feinberg, 2010).

4.2.2.4 Procedure

To meet both the COVID-19 restraints measures and the professionals' agendas, the interviews were carried out via a teleconferencing system at day and time scheduled with participants. On the day of the interview, the participant was welcomed, then the researcher introduced himself and presented the project. Then the interview started following the script (section 2.3). During the interview, the participant was shown the video (prompt) introducing the topic of 3D web (in the case of retailers) or AR (for trade fairs operators). At the end of the interview, the researcher shared the link to answer a short questionnaire about the technological experience (section 2.2.1). The interviews lasted at most about one hour. To allow subsequent data analysis, the interviews were video recorded. The recordings (approximately 11 hours) were then transcribed verbatim, for a total of 125 pages analyzed (Halcomb & Davidson, 2006); specifically, the researchers initially read the transcripts carefully noting possible emerging themes, and subsequently, they compared the identified themes and agreed on shared categories and their titles.

The present research was designed in accordance with the Declaration of Helsinki and approved by the ethics committee for psychological research at the University of Padova.

4.2.3 Results

We first report the findings from the technology experience questionnaire (Table 4.1). Subsequently, we report the findings emerging from the interviews run with retailers. Then, results from the interviews with fair trade operators are presented.

	M	SD
Personal innovativeness	3.92	.865
Technology anxiety	4.04	.797
Perceived relative advantages	3.59	.712
Perceived compatibility	3.94	.479
Perceived complexity	3.41	.554
Perceived trialability	4.41	.861

Table 4.1: Technology Experience Questionnaire (mean and standard deviation).

4.2.3.1 Retailers' experience with e-retail

With regard to 3D web technology, the four main topics explored were: the experience and issues with e-commerce, e-commerce features and management, time required to upload and update product-related information and the importance of direct contact with customers.

Experience and issues with e-commerce. Half of the retailers interviewed (N=6) owned and still use their own website for e-commerce, while one retailer (N=1), in addition to having its own e-commerce, also employs services by third parties (e.g., Subito.it, eBay). These retailers trade in furniture, bookshops, jewelry, food retailing, hi-fi, clothing, electronics, and household appliances. One retailer (N=1) working in the gastronomy sector reported not to have a proprietary e-commerce but relies on third-party service to deliver products. On the contrary, 34% of the retailers (N=4) reported that they had a negative experience in the management of their e-commerce sites. As a result, they closed the online shop. In detail, these retailers deal with flowers, ironware, and beauty. The main issues that emerged were different from case to case. For the flower shop, the main problem was the difficulty in meeting the customer's expectations due to the differences between the images published online and the physical product (especially the size). The ironware shop reported that the main concerns were logistical issues in warehouse management and shipment. In the case of the beauty, the main issues were the poor interaction with customers and the difficulty to regularly update it. Lastly, in the case of the jewelry shop, the issue amounted to a highly negative experience with the company responsible for managing the website.

Almost half of the e-commerce owners (N=3) reported that they had a website for the purpose of showcasing and not selling. In addition, all retailers (N=12) are on social media such as Instagram and Facebook, in an attempt to establish and maintain a direct interaction with their customers and to increase their visibility.

E-commerce features and management. With the aim of understanding which attributes of the product are more relevant when uploading a product online, retailers were asked what attributes they entered for their products. What can be observed is that all retailers emphasize the importance of the descriptive and visual aspects (images) of the product. Some focus only on these two aspects, others report that they also focus on aspects such as uploading videos or video tutorials, conveying the values of the company, the price, and the availability of additional services. The answers collected related to participants' current or past experience with their e-commerce website.

Moreover, four respondents (N=4) reported that they also uploaded emotional contents, such as images or videos. Three retailers (N=3) reported that they include (or, in one case, intended to include) video tutorials in the attempt to provide a better description of their products. One trader (N=1) reported inserting a textual description of the product with the purpose to index some keywords on search engines. Finally, two retailers (N=2) reported that the type of information varies according to the type of product on sale, and which of the attributes they decide to pay more attention to (especially in the case of the bookshop).

Time required for uploading and updating product-related information. Only two retailers (N=2) managed autonomously the upload of product-related information, and they reported that this activity takes them a little time. Regarding the frequency of updates, four (N=4) reported that they spend approximately a few hours per week on this activity. Only one participant (N=1) reported dedicating a few hours monthly to this activity.

Direct contact with the customer. Overall, 82% (N=9) of retailers reported that they consider it necessary for the customer to physically visit the shop. Indeed, only by visiting the shop, the customer can try the products (i.e., hi-fi, clothing, floristry) or can receive full advice on the product. On the contrary, 18% of retailers (N = 2) do not consider physical contact with the customer to be necessary (i.e., ironware and food shops).

In addition, all retailers (N=12) have stated that they also use social media in order to maintain contact with their customers or to create a new direct sales channel. In fact, some of them reported that the use of social media has been important during the lockdown periods because it allowed direct communication with customers for issues related to product shipment and sales advice. After watching the video illustrating the 3D web, the majority of respondents (N=7) could not contextualize the technology for their shops. In two cases (N=2) retailers reported distrust towards this technology. More generally, respondents expressed concern on the management of the technology.

4.2.3.2 Trade fairs operators' experiences

Respondents reported that today the main functions of the fair continue to be to establish and maintain contacts with their clients, and to keep updated on the trends in a specific sector.

Participants reported that the COVID-19 has changed the shape of trade fairs. Indeed, they have been attending digital trade shows for two years.

Managing the stand. Overall, 75% of the exhibitors (N = 3) stated that they use modular stands, which are convenient because they can be adapted according to the space allocated by the event. On the other hand, the remaining exhibitor reported using stands made ad hoc by third parties, which are easy to replicate in the different events attended by the company. Moreover, exhibitors reported to use several technological devices such as computers, monitors, and backlit panels to better display the products. According to respondents, it takes between one day and one week to set up these complex structures.

Exhibitors were asked whether the current arrangement of the stand allowed them to reach all interested visitors. In general, exhibitors (N=3) feel that their display method allows them to reach all interested parties, although this requires additional work to schedule purposeful visits.

Selection of the products. Exhibitors were asked about their strategy for selecting the products to be shown at the fair. In general, from the interviews it emerged that it is not possible to exhibit in the stand all the products and, therefore, it is fundamental to make a careful selection. More specifically, the respondents reported that the basic strategy is prioritizing the top-of-the-range products (regardless of the selling volumes). Additional drivers for the choice of the products are the market trends of the country where the fair will be held and the characteristics of the customers attending the event. However, as mentioned above, emerges the importance along with the difficulty, of being able to show as many products as possible.

Integration of AR. In general, four operators (N=4) and the stand designer were interested in the adoption of AR for product display at the fair. Still, they all reckoned that the technology needed in-depth studies to understand how to fully exploit it and make it meaningful for their products. In any case, they believe that AR could give them the possibility to display a greater number of products at low costs, avoiding transport costs and the risks it entails.

One of the exhibitors (N=1) even suggests that this technology could be useful in illustrating product components in more detail. In addition, the designer expressed interest in applying augmented reality in the context of booth design as well, particularly in reference to product measurements to be included in the booths and the graphic aspect.

Although the impressions are generally positive, for the two remaining operators (N=2), the application of AR as an exhibition medium is not interesting because they consider it still immature and not equal to physical reality. While for the designer, the problem linked to the adoption of AR is due to the fact that in most cases company managers are lagging behind technological evolution.

4.2.4 Discussion

The present work investigated the factors supporting and hindering the adoption of e-commerce by small businesses. Furthermore, it addressed how professionals think 3D web could be integrated into their businesses. The findings emerged from the interviews confirmed the obstacles identified by previous research, mainly concerning organizational and technological factors (Hong & Zhu, 2006; Wymer & Regan, 2005). We also translated those findings into operative guidelines that can inform the design of interfaces meant to enable small businesses to launch and manage their own e-commerce. The guidelines emphasize the importance of an easy and effortless information input and management to smoothly integrate with all the other work activities retailers have to carry out. Additionally, we investigated how operators in the sector of trade fairs consider the application of AR technology. To the best of our knowledge, this is the first work addressing this topic. Interestingly, we found that professionals were generally open to the introduction of AR, because it would allow them to bring to the events a significantly larger number of products without increasing the costs related to the organization and transportation. Moreover, it opens possibilities to explore the products from novel perspectives. Still, it is reckoned that AR could bring an added value to trade fairs when it is employed as an addition to the traditional stand, and not as a replacement to it.

4.2.5 Design Guidelines

The findings from the interviews were translated into design guidelines, meant to inform the design of the interfaces to be developed within the e-shop.

First, we report on the user requirements for 3D web interfaces extracted from the interviews with retailers:

- Intuitive content upload: the interface should be simple and immediately understandable interface, to allow retailers to manage the shop autonomously and with little effort;
- Multi-modal product information input: the interface should support the upload of various media (technical description, emotional description, images, emotional photographs, video clips, video tutorials) to describe the products;
- Modularity management of product-related information: the interface should enable retailers to emphasize specific pieces of information;
- Customer care: the interface should feature a function to establish and maintain direct contact with the customers.

Secondly, guidelines for the design of augmented reality interfaces for the trade fair context are presented below, extracted from the investigation of trade fairs operators. In detail:

- Seamless integration: currently, trade fairs are composed of content and events that take place in-person and some content that is accessible online. Augmented content should integrate with the exhibition at the stand, not replace it;
- AR catalog: thanks to AR, exhibitors should be able to demonstrate to visitors all the products that they cannot display on their stand, for reasons of space, logistics and costs. In other words, AR should allow exhibitors to have a virtual catalog of their products that can be explored in full size;
- Contextualization of product features: products offered in AR should highlight product features that are normally not very visible;
- Customization: unlike the products physically displayed on the stand, products in AR must be able to be customized according to the customer's needs, compatible with the characteristics of the product.

4.2.6 Conclusion

The impact of online retail is ever increasing, and the restrictions brought about by COVID-19 have even emphasized this trend (Jílková & Králová, 2021). Still, small businesses struggle to keep up by launching and maintaining their own shop online. The B2B sector was also seriously damaged by the pandemic, with all the events and exhibitions suspended for more than two years. In the present work, we describe the barriers encountered by small retailers in administering their own e-commerce. Additionally, we also report the design guidelines for the developments of e-commerce platforms addressed to small-sized businesses, thereby extending previous research. We further considered the deployment of AR technology in the trade fair environment, finding a substantial openness.

4.3 Customer experience with 3D virtual objects in AR shopping

4.3.1 Background

4.3.1.1 The potential of AR in shopping experience

Over time, augmented reality technology has assumed an important role in various fields such as medical treatment and surgery, tourism, cultural heritage, manufacturing, automotive industry, education, and marketing (Chen et al., 2019; Dargan et al., 2022). One of the most promising applications of AR is in the shopping experience, where it can provide a more immersive and interactive environment for consumers to browse and purchase products.

Specifically, recent studies have shown that augmented reality shopping experiences have a positive effect on the consumer, for example, by improving the flow experience in users through a more vivid graphical representation of the object and also by giving them more informative details (Barhorst et al, 2021). Moreover, it has been shown that exposure to 3D virtual objects improves the mental fluency of the product image, in turn increasing purchase intention and attitude toward the brand (Hilken et al., 2021). Also, Escobar et al. (2021), in the context of food service, showed that the mental simulation of the act of eating is greater in consumers when a product is shown in augmented reality rather than through images. In general, augmented reality influences several aspects of the users' shopping experience by modifying, more or less directly, their behavior. For instance, the interaction with a product, makes users feel more confident in their choice thanks to a more informative and playful experience (Kang et al., 2020). Also, the possibility to customize a product, by changing its shape and color, is another feature that increases the user's psychological ownership of the product (Jussila et al., 2015).

Most e-commerce systems using augmented reality today leverage mobile applications that integrate virtual objects into the user's real context, projecting the virtual object into real space through the use of a screen (hand-held devices). This is the case with the Amazon and IKEA e-commerce applications mentioned above. These systems allow users to visualize the product in the real environment, but do not allow them to manipulate it freely. In fact, the only functions that users can control are positioning and horizontal rotation on the product's vertical axis.

4.3.1.2 Interaction with Virtual 3D Objects

Gestures are recognized as a significant modality of expression and communication. They are considered fundamental and innate kinesics components, comprising body postures, movements, and expressions that convey information that aligns with our discourse (Muller et al., 2017). For this reason, during Augmented Reality experiences, researchers have investigated gestures that enable users to explore virtual objects in the most natural way possible, thereby combining the intuitiveness of exploring physical objects with the additional features of Augmented Reality (Ortega et al., 2017). This involves identifying gestures that are similar to those used to manipulate physical objects, such as grasping and rotating, as well as those unique to Augmented Reality, such as zooming and scaling (Hertel et al., 2021). Therefore, identifying the most natural interaction can become a complex process. This is due to the variability of our interaction with objects, depending on the context, which consequently demands an in-depth understanding of our ways of interacting with physical objects in our daily lives (Kang, 2020).

Recent studies (Pham et al., 2018; Ortega et al., 2017) investigated the relationship between the size of virtual objects and the corresponding gestures used to manipulate them. Pham et al (2018) discovered that people tend to interact with virtual objects taking into account their physical features. Specifically, people appear to be influenced by the affordances of objects, adjusting their gestures in response to variations in object size, distance, and scene size. As an example, the size of the virtual object has an impact on the number of hands used to manipulate it. Smaller-size objects are typically grasped with a few fingers, whereas larger-size objects are typically grasped with one hand until they reach a size that requires the use of both arms, as if the virtual object was perceived to be heavier. Similarly, the distance between the user and the object affects the type of gesture used: people tend to use proximal gestures and appear to touch small-size objects, while larger-size objects are selected from a distance using distal gestures. These findings suggest that our interaction with virtual objects is influenced by their physical features, and that our behavior when interacting with virtual objects is modeled after our behavior when interacting with their physical counterparts. Moreover, other studies (Plank et al., 2017; Tarre et al., 2018) suggest that users tend to favor wider actions when interacting with larger objects as opposed to smaller ones, even when performing the same action.

The manipulation of virtual objects has also been studied in immersive virtual reality environments: several studies have examined whether the size of virtual objects has an impact on the mid-air manipulation process. Empirical findings have indicated that the performance of manipulation tasks decreases significantly if virtual objects are too small or too large (Viola et a., 2022; Gloumeau et al., 2020) and that the time required to select small objects is typically twice as long as the time needed for larger objects (Wang et al., 2021). Consequently, these issues lead to a less natural and more challenging interaction, considering both Virtual and Augmented contexts (Englmeier et al., 2020; Zhenyi He 2014).

To the best of our knowledge, despite the substantial body of research on eliciting gestures in Augmented Reality (Piumsomboon et al., 2013; F. R. Ortega et al., 2017), there remains a significant gap in knowledge regarding whether the size of virtual objects affects manipulation preferences in AR.

4.3.2 Data collection

The aim of the present data collection was to explore whether the size of virtual objects affected their manipulation modality, by comparing manual free-hand gestures with automatic scaling via anchoring, which allowed users to view the product at 1:1 scale.

Given the success of mobile Augmented Reality apps which allow for the automatic scaling of selected products such as large furniture to match the dimension of a shopper's room (Ozturkcan, 2021), it is conceivable that users would also prefer this feature when using head-mounted displays in Augmented Reality environments. For this reason, when manipulating large-sized objects, we hypnotized that users could find more attractive and efficient an "automatic scaling by anchoring" interaction mode (H1). Differently, for small objects we expected users' preferences of manipulations that allows them to explore the object by touching and rotating it, as is the case with physical products. Consequently, our hypothesis was that participants, when manipulating small objects, displayed a preference for the "free manipulation" interaction modalities (H2).

In other words, we excepted a better user experience in the "free size manipulation" condition when compared to the "automatic resize by anchoring" condition for small products. Conversely, we hypothesized that user experience would be better in the "automatic scaling by anchoring" condition than in the "free manipulation" condition, for what concerns large products.

4.3.2.1 Experimental Design

A 2 (manipulation of the objects) x 2 (size of the objects) within-participants design was employed. The two levels of the first variable were free scaling manipulation and automatic scaling by anchoring, while the two levels of the second variable were small-sized products and large-sized products.

The free scaling manipulation allows participants to interact with 3D virtual products (i.e., small-sized and large-sized) through manual manipulation, scaling objects without necessarily respecting their actual size. Therefore, He/she could freely manipulate the products by pinching, moving, rotating, enlarging, and reducing them according to his/her preferences. The concept of "scaling" refers to the geometric manipulation technique (Hertel et al., 2021) employed to reduce or enlarge the size of an interactive element (Spittle et al., 2022).

Differently, the automatic scaling by anchoring allows participants to manipulate the 3D virtual products (i.e., small-sized and large-sized) scaling the objects to their effective size. Users could manipulate the objects by grabbing them using the pinch action and moving them to the sphere-shaped anchor points (present to the left of the exhibitor) which allowed viewing the product on a 1:1 scale.

4.3.2.2 Equipment and technical set up

HoloLens 2 and gesture

The participants were provided with Microsoft HoloLens 2nd generation smart glasses (Windows Holographic OS, SoC Qualcomm Snapdragon 850 2.95 GHz, custom-built Holographic Processing Unit 2.0, 4GB LPDDR4x system DRAM, see-through holographic lenses 2K 3:2, 43° horizontal field of view, weight 56g). As a fully untethered holographic computer, neither cables nor further devices were needed for executing the experiment. The augmented environment was programmed ad-hoc in Unity (version 2020.1.17f1) for the present investigation, instead, the 3D models of the objects have been downloaded from various online databases (i.e., Sketchfab, CGTrader, TurboSquid), under free or paid license, and were then modified using Blender to be compatible with the Augmented Reality application. Participants interacted with the augmented objects by physical collision. The gestures performed by the participants were the following:

- *Grab*: by placing an open hand in front of the body, a ray appears to indicate the direction of the interaction. By directing this ray at an object and closing the thumb and forefinger (pinch gesture) the object is grabbed. At this point, while maintaining the pinch, it is possible to move the object by moving the hand.
- *Object scaling*: to resize an object, it is necessary to use the 'pinch' action on it with both hands at the same time. At this point, you only need to widen or narrow the position of your hands to enlarge or shrink the object.
- *Object rotation*: by holding the object using the pinch action with one hand and rotating the wrist, it is possible to rotate the object in any direction. Alternatively, it is also possible to rotate the object using two hands by performing the pinch action on the object with both hands at the same time (similarly to the scaling gesture) and rotating them in the desired direction.

Virtual 3D products

The experiment involved two categories of augmented products: small products and large products.

Products are composed respectively of two sub-groups of products: household appliances and home goods presented in a physical exhibitor. The two sub-group of products allowed participants to interact with different products under all experimental conditions.
As far as small objects are concerned, the group of home goods was made up of clocks, lamps, lanterns, and vases; while the group of household appliances was made up of boilers, coffee machines, planetary mixers, and toasters. As far as large objects are concerned, home goods is characterized by armchairs, chairs, shelves, and tables; while household appliances was made up of washing machines, smart TV, fridges, and ovens. For each product, two different models have been proposed, for a total of 32 different models.

Setting and physical exhibitor

Participants were tested indoors in a room on the local university campus with constant neutral lighting. The field of action was about 3 × 3m, whereby participants were free to walk and interact with the AR stimuli. The 3D virtual objects were presented in a physical exhibitor composed of 16 compartments. The position of the exhibitor was recognized by the headset through the object recognition system enabled by the Vuforia Engine software development kit (SDK).
Virtual products have been resized (all to the same size) to fit in a 20-centimeter cube and placed in the compartments most visible to users (i.e., upper empty compartments; Figure 4.1).

Figure 4.1 Physical exhibitor with exposed 3D virtual products. Specifically, small furniture products are displayed on the left, while large furniture products are displayed on the right.

4.3.2.3 Purchasing scenarios

Participants were asked to explore two products at time and purchase the one that met specific requirements. Task scenarios were devised to help participants better contextualize the shopping experience. More specifically, the participants had to take the two products in question off the shelf and explore them carefully, comparing them, in order to buy the right product (without limits of time or money). Once they had decided on the product to purchase, they had to put it in the shopping cart and replace the unselected product on the shelf. An example of a proposed

purchasing scenario is shown below: "*You have been asked to buy a new coffee machine for the office. Your company is particularly sensitive to environmental issues, so great care is always taken in making environmentally friendly choices. Identify the coffee machine that allows you to use both compostable pods and coffee powder and thus not pollute, and buy it by putting it in your shopping cart*".
Participants were explained how to interact with objects according to the experimental condition.

4.3.2.4 Measurements and materials

Behavioral data

Number of manipulations. The number of manipulations performed by the participant was tracked for each individual object (each manipulation began when an object was grabbed and ended when the object was released).

Duration of manipulations. The duration of manipulations performed by the participant was tracked for each individual object.

Time to complete each experimental condition. We measured the execution times of each experimental condition (start: object loading; end: placement of the 3D virtual object in the shopping cart).

Purchasing success. the percentage of correct and incorrect and missed trials purchased product as a measure of task accuracy.

Behavioral data were automatically saved as a data-log file on the HoloLens internal storage (64 GB Universal Flash Storage 2.1) at the end of each experimental session.

Self-report data

Socio-demographic data (age, gender, educational level) and information about previous experiences with augmented reality viewers and with the HoloLens 2 headset were collected to understand users' backgrounds. Indeed, online purchasing habits (on e-commerce such as Amazon, applications such as Glovo, etc.) were investigated for specific product categories: electronics and household appliances, food and beverages, cosmetics and personal care products, clothing and accessories, and components of furniture.

Need for touch (NFT). The need for touch is conceptually defined as a preference for obtaining and using information through the haptic system (Peck & Childers, 2003). This is a multidimensional construct, with two main aspects: instrumental and autotelic. The instrumental dimension of NFT refers to those aspects of pre-purchase touch that reflect outcome-directed touch with a salient purchase goal. Autotelic touch, on the other hand, is touch for the sake of enjoyment and sensory stimulation, without necessarily having a specific purchase goal in mind;

it involved a hedonic-oriented response (Peck & Childesrs, 2003; Holbrook & Hirschman 1982). Overall, the NFT is a concept that highlights the importance of touch as a sensory modality in consumer behavior. Measuring individual differences in NFT, allow us to better understand how touch influences consumer decision-making and product evaluation.

This construct was measured using the Need for touch scale (Peck & Childers, 2003). The scale consists of 12 items in total, 6 of which explore the autotelic dimension and 6 of the instrumental dimensions during a traditional shopping experience (e.g., supermarkets, malls, clothing stores, etc.). Participants responded on a 7-point Likert response scale (1 = strongly disagree; 7 = strongly agree).

User experience (UX) with object manipulations. Marc Hassenzahl's (2010) model of the user experience conceptualizes it in terms of pragmatic and hedonic aspects. By pragmatic, it is meant how simple, practical, and obvious it is for the user to achieve their goals. By hedonic, it is meant how evocative and stimulating the interaction is to them.

The UX of manipulating 3D virtual objects was evaluated through the User Experience Questionnaire. It is composed of 6 scales with 26 items in total (Schrepp, 2015). *Attractiveness* was a pure valence dimension and investigated the overall impression of the manipulation modality (6 items). *Perspicuity* assessed whether it was easy to become familiar with the manipulation modality and whether it was easy to learn how to use it; *efficiency* explored whether users could solve their tasks without unnecessary effort; *dependability* investigated if the user feels in control of the interaction; *stimulation* investigated if it was exciting and motivating the manipulation modality; n*ovelty*: assessed whether the manipulation modality was innovative and creative and if catch the interest of users. *Perspicuity*, *efficiency,* and *dependability* are pragmatic quality aspects (goal-directed), while *stimulation* and *novelty* are hedonic quality aspects (not goal-directed); each scale consists of 4 items. The items have the form of a semantic differential, i.e. each item is represented by two terms with opposite meanings placed at the extremes of a 7-point scale.

Exciting quality. Exciting quality refers to the degree to which an experience elicits a combination of pleasant and arousing (adapted from Russell & Pratt, 1980). It provides a useful framework for understanding how individuals respond to different object manipulations and can help inform the design and development of more engaging and effective AR systems. The exciting quality of object manipulations was investigated through 5 items adapted from the exciting quality scale by Russell & Pratt (1980). Items are provided by a semantic differential on a series of bipolar adjectives (e.g., exciting-dull) placed at the extremes of a 7-point scale.

Presence of 3D Virtual models in AR environment. It measured the experienced presence of virtual objects in the physical environment (Regenbrecht & Schubert, 2021). Thus, it measures the level of perception that VOs are present in physical space, although not identical to it, and that they achieve a certain degree of realism (Banos et al., 2000). This construct was investigated through the presence in an augmented environment questionnaire by Regenbrecht & Schubert (2021). It was composed of 7 items that explored three different dimensions: *realness*, *spatial presence,* and *perceptual stress*. *Realness* combined items related to how real the virtual objects seemed, and how well they integrated with the physical objects, while *spatial presence* combined items on the 3Dness of the virtual objects. *Perceptual stress* explored whether the difference between physical and virtual attracts attention and whether the perception of the virtual object requires effort to be placed in the physical environment. Participants responded on a 7-point Likert scale (1= totally disagree; 7= totally agree).

In addition, to evaluate users' shopping experience in AR shops, a number of aspects were investigated below.

Quality of 3D Virtual models. It had the objective of evaluating the quality of the 3D virtual models in terms of graphic satisfaction and vividness of rendering. To do this, two items ad-hoc were realized. Particularly, for each product explored during the experience (represented by an image to facilitate participants) participants provided a score from 1 to 10 with respect to their graphic satisfaction and vividness.

Customer satisfaction. To evaluate customer satisfaction related to AR shops and products, a number of aspects were investigated below.

Overall satisfaction was evaluated through one item (5-point Likert scale) adapted from Poushneh and Vasquez-Parraga (2017; Taylor & Baker, 1994). *Desire to stay* is an affective state of attraction to the environment that motivates the customer to remain in the environment for an extended period (Wakefield & Blodgett, 1994). It refers to the degree to which customers feel comfortable and engaged in the physical space and atmosphere of the service setting. To measure the degree to which customers feel engaged in the shop, we adapted one item from Wakefield & Blodgett (1994) on a 5-point Likert scale. *Repatronage* or purchase intention is a concept that refers to a customer's likelihood of purchasing a product or taking benefit of the service in the future (Oliver & Swan, 1989). We evaluated the intention to purchase again in the augmented reality shop through 4 items adapted from Oliver and Swan (1989). Items are provided by a semantic differential on a series of bipolar adjectives placed at the extremes of a 7-point scale. In addition, we measured premiumness expectations by adapting the *luxury dimension* (Ko et al, 2019) as in the previous studies (Barbosa Escobar et al., 2021). Specifically, 4 items assessed product perception in terms of quality, authenticity, willingness to pay a premium price, and also general premiumness (5-point Likert scale). Finally, *quality concern of physical product* was investigated by 2 items (adapted from Kuhn, Lichters & Krey, 2020; Imschloss & Kuehnl, 2017;

Krishna & Morrin, 2008) to understand if the quality of the products that will arrive at home was perceived worse/better than it appears (10-point scale).

Overall user preferences. Once the shopping experience proposed was finished, users' preferences with respect to manipulations of 3D small and large products in the AR environment were assessed. To do this, a post-experience questionnaire was carried out considering *ease of use*, *effectiveness, utility, and pleasantness* through 8 *ad hoc* items (4 related to large objects and 4 related to small objects). The items have the form of a semantic differential, i.e., the two different manipulations were placed at the extremes of a 7-point scale in order to understand users' preferences.

Wearable device technology acceptance. Acceptance of wearable technology refers to the willingness of individuals to adopt and use wearable devices in their daily lives. We investigated this construct with particular reference to the head-mounted display for shopping in AR shops. To do this, we adapted the acceptance of wearable technologies questionnaire from Spagnolli and colleagues (2014). The questionnaire was intended to investigate the general attitude towards technology and, more specifically, is adapted to the opinion on the AR device worn during the experience. The 26-item (6-point Likert response) was composed of the following dimensions: *Attitude toward technology* (6 items): regards an individual's overall reaction toward the use of technological devices; *technology anxiety* (5 items): refers to the feeling of apprehension or even fear when using technology*; facilitating conditions* (3 items): refers to the extent to which the user perceives that there are factors able to support the system deployment, *perceived usefulness* (4 items): refers to the extent to which the user believes that the use of a certain device would enhance her performance, *effort expectancy* (2 items): refers to the perceived ease associated with the use of a certain device, *behavioral intentions* (4 items): refers to the extent to which an individual has formulated conscious plans to carry out a certain action using a particular device; *psychological attachment* (5 items) dimension refers to the individual's tendency to adopt a technology as a consequence of social influence; *perceived enjoyment* (4 items) refers to the extent to which users perceive the activity of using a system to be pleasant, aside from any consequences resulting from device use; *perceived comfort* (7 items): refers to individuals' perceived comfort in wearing a wearable computer; *perceived privacy* (5 items): refers to the extent to which the user is confident that the data recorded by the device is safely handled and stored.

4.3.2.5 Procedure

Pre-experimental phase

After signing informed consent, all participants filled out a demographic questionnaire and Need for touch scale. Then, a pre-training session and a training session were held, preparatory to carrying out the experimental session.

Pre-training session. The pre-training session was held aimed at familiarizing oneself with augmented reality technology and the HoloLens 2 device. Once the headset was worn, the participant was asked to stand in front of the exhibitor and read the text instruction (also supported by videos) relating to the gestures to be performed to manipulate virtual objects. Subsequently, the participants performed the gestures and interacted with two 3D virtual geometric objects (i.e., a pyramid and a cube) through 8 tasks provided verbally by the experimenter (e.g., task "*Now I ask you to grab the withe object and place it on the exhibitor in front of you*"). This phase required a maximum execution time of 10 minutes.

Training session. The training session in the augmented environment was based on the same tasks used afterward in the experimental phase; therefore, it envisaged the use of the physical exhibitor. The aim of this practice was to familiarize with the augmented stimuli and minimize the incidence of individual differences in the ability to use the experimental apparatus. The participant was invited to position himself in front of the exhibitor, which displayed a series of 3D virtual geometric objects. The instructions for interacting with the virtual elements were presented through the AR system as in the pre-training session, and the tasks to be performed were provided by the experimenter (e.g., task "*Now I ask you to extract the red ball and enlarge it so that it has a diameter of about one meter, then I ask you to make it smaller and put it back in its place on the shelf.*"). The researcher monitored the participant from the desktop and assigned a score to each sub-task on a previously created *ad hoc* grid. If the subject completed the pre-training in a maximum of 10 minutes and the training with a score of 100%, the participant was moved to the experimental phase, otherwise, the participant was thanked and dismissed.

Experimental phase

The participant was introduced to a purchase scenario. On the shelf were 8 virtual 3D products for sale and to the right was a virtual shopping cart. The participant was instructed to take two products of the same sub-group (related to the scenario described off the exhibitor and explore them carefully, comparing them, in order to purchase the right product). Once the participant had decided on the product to purchase, he had to add it to the cart and put the unselected product back on the shelf. Participants repeated the task 16 times in total, 4 for each experimental condition: small furniture products + free manipulation modalities (session 1), large furniture products + free manipulation modalities (session 2), small domestic appliances + anchor manipulation modalities (session 3), large domestic appliances + anchor manipulation modalities (session 4). The order of presentation of the 4 experimental conditions was randomized (Figure 4.2).

After each experimental session, participants filled in the questionnaires to measure: UX with object manipulation, excitement quality, presence in AR environment, quality of 3D virtual models, and customer satisfaction. Finally, at the end of the data collection, questionnaires were administered to investigate overall user preferences and wearable device technology acceptance. Self-reported data were collected through the use of Qualtrics[12XM] software. The total duration of the data collection was about 60 minutes.

The present research was designed in accordance with the Declaration of Helsinki and approved by the ethics committee for psychological research at the University of Padova.

Figure 4.2: An example of exploring a small product through the free manipulation modality (left) and an example of exploring a large product through the automatic manipulation modality (right).

4.3.2.6 Participants

A total of 40 participants volunteered to take part in the data collection (without compensation). An a priori power analysis conducted on Gpower (Erdfelder et al., 1996) indicated that a total sample of 40 people was needed to detect a medium effect size (d = .5) with 95% power. All participants provided informed consent prior to engaging in the data collection.

The inclusion criteria were: (a) between the ages of 18 and 40, (b) excellent comprehension of the Italian language (c) no significant vision problems and no use of glasses or contact lenses; (d) absence of psychopharmacological treatments. The participants were 40 young adults (20 female), with ages between 21-33 (M_{age}= 26, SD_{age}= 3.10) and an average level of education ($M_{education}$= 15.65, $SD_{education}$= 2.70, $min_{ducation}$= 8, $max_{education}$= 21). None of the participants had current or past neurological or psychiatric problems. All the participants had normal or corrected-to-normal visual acuity and reported having a normal color vision.

[12] https://www.qualtrics.com

Most of the sample (N= 29, 72.50%) had never used an AR headset, only 5 participants have used it one time (12,50%) and 5 participants few times (12,50%), and only one participant declared to use it often (2,50%).

Of the 11 participants who had used an augmented reality headset before, 1 participant used HoloLens often, 1 participant sometimes, and 4 participants only once, while the remaining 5 participants did not know what type of viewer they had used. In addition, participants provided information about their online shopping frequency for different product categories: electronics and household appliances, clothing and accessories, and cosmetics and personal care products are purchased more frequently online, unlike food and beverages and components of furniture (Table 4.2).

Online purchase frequency for different product categories					
Descriptive Statistics	electronics and household appliances	food and beverages	cosmetics and personal care products	clothing and accessories	components of furniture
M	2.95	1.85	2.23	2.75	1.85
SD	.82	1.25	.97	.81	.98

Table 4.2: Descriptive statistics about the frequency of online purchases of different product categories:1= never, 2= rarely (less than once a month), 3= sometimes (once or twice a month), 4= often (once a week), 5= regularly (every day or nearly so).

The sample involved showed good levels of acceptance toward the augmented reality viewer in the context of shopping. The technology acceptance questionnaire was analyzed by performing a series of Wilcoxon tests to compare the score provided by participants with the median response scale (Mdn = 3). The results were then corrected by the B.H. method (Benjamini & Hochberg, 1995). The dimensions which resulted significantly higher than the median of the scale were: Attitude Toward Technology (V = 820.00, p < .001; Mdn = 5.00, M= 4.85, SD = 0.74), Technology Anxiety (V = 720.00, p < .001; Mdn = 4.67, M= 4.45, SD = 1.10), Facilitating Conditions (V = 439.00, p< .001; Mdn = 4.00, M = 4.28, SD = 1.06), Effort Expectancy (V= 715.00, p < .001; Mdn = 5.00, M = 4.88, SD = .94), Behavioral Intention (V= 692.00, p < .001; Mdn = 4.67, M = 4.43, SD = 1.23), Perceived Privacy (V = 297.00, p = .03; Mdn = 3.50, M= 3.81, SD = .88), Perceived Comfort (V = 765.50, p < .001; Mdn = 4.59, M= 4.65, SD = .79). Differently, they exhibited neutrally regarding Perceived Usefulness (V= 434.000, p = .75; Mdn= 3.670, M= 3.59, SD= 1.09) and Psychological Attachment (V= 261.00, p = .11; Mdn= 3.000, M 3.14, SD= 1.34).

4.3.3 Analysis and Results

This section reports the analyses conducted with the R studio software and the results obtained.

4.3.3.1 self-report data

Questionnaires filled at the end of each experimental session are conducted are analyzed as follows.

A series of Shapiro-Wilk tests highlighted the non-normal distribution of the data. Therefore, non-parametric analyses were performed.

A series of Wilcoxon tests were conducted to determine if there were any effects on the dependent variables measured for the size of the object, the manipulation of the object, and the interaction between the two variables (object size - object manipulation). The p-values obtained from the various analyses were adjusted with the application of the B.H. method (Benjamini & Hochberg, 1995). The following is a detailed explanation of the data analyzed:

- Object size variable. The scores given by manipulating small objects (mean of scores related to data obtained from free scaling and automatic scaling of small objects) were compared with the scores given by manipulating large objects (mean of scores related to free scaling and automatic scaling of large objects).
- Object manipulation variable. The scores obtained during the free scaling of the object (mean of scores related to data obtained from free scaling of small objects and large objects) were compared with scores obtained during automatic scaling (mean of scores related to automatic manipulation of large objects and from automatic manipulation of small objects).
- Interaction (object size - object manipulation). One delta was calculated by subtracting the data obtained during the free scaling of large objects from the data obtained during the automatic scaling of large objects. The other delta was calculated by subtracting the data obtained during the free scaling of small objects from the automatic scaling of small objects.

In cases where statistically significant results emerged, a series of post-hoc tests were performed with the application of B.H. correction for multiple comparisons. Specifically, comparisons with Wilcoxon signed-rank tests for paired samples (B.H. correction) were conducted.

Finally, for each dimension investigated through the questionnaires, a series of Wilcoxon tests (B.H. correction) were conducted in comparison with the median of the response scale in order to verify whether the opinions provided by the participants were positive or negative.

Users' opinions on the manipulations of virtual objects

The table below shows the results of the Wilcoxon tests that emerged for each dimension of the UEQ (Table 4.3).

User Experience Questionnaire				
Dimensions	Indipendent variable	V	z	p (B.H.)
Attractiveness *	Object Size	518.00	2.91	.01
	Object Manipulation	394.00	.96	.34
	Object Size x Object Manipulation	522.50	3.40	<0.01
Perspiquity *	Object Size	393.00	1.63	.10
	Object Manipulation	467.00	2.90	.01
	Object Size x Object Manipulation	447.00	1.79	.10
Efficiency *	Object Size	540.50	3.26	<0.01
	Object Manipulation	568.00	2.86	.01
	Object Size x Object Manipulation	483.00	1.98	.04
Dependability	Object Size	382.50	.47	.64
	Object Manipulation	451.50	1.17	.37
	Object Size x Object Manipulation	450.00	1.15	.37
Stimulation *	Object Size	323.50	1.11	.40
	Object Manipulation	303.50	.41	.69
	Object Size x Object Manipulation	389.50	3.23	<.01
Novelty	Object Size	274.00	-0.67	.51
	Object Manipulation	278.00	-0.86	.51
	Object Size x Object Manipulation	357.50	.70	.51

*Table 4.3: Results of Wilcoxon tests on user experience questionnaire; (*p(B.H.) < .05).*

Post-hoc analyses showed that it is more attractive (V = 461, p = .001) to manipulate large objects (Mdn = 2.00, M = 1.76, SD = 0.98) than small objects (Mdn = 1.17, M = 0.97, SD = 1.19) by automatic scaling. Moreover, it is more attractive (V = 515.50, p < .01) to manipulate large objects by automatic manipulation (Mdn = 2.00, M = 1.76, SD = 0.98) than by free one (Mdn = 1.12, M = 1.21, DS = 1.00). Moreover, greater perspicuity emerged (V = 495.50, p < .001) for manipulating large objects with automatic resizing (Mdn = 2.625, M = 2.28, DS = .79) than for manipulating objects of the same size with free resizing (Mdn = 1.75, M = 1.40, DS = 1.26), and then for small objects with both automatic (V= 337.50, p = 0.02; Mdn = 2.00, M = 1.83, DS = .94) and free resizing (V = 425.00, p < .001; Mdn = 1.62, M = 1.51, DS =1.06). In addition, small objects with automatic manipulation (M = 1.83, DS = .94) have greater perspicuity than objects with both small (V = 466.00, p < .05; Mdn = 1.62, M = 1.51, DS =1.06) and large (V = 352.50, p < .05; Mdn = 1.75, M = 1.40, DS = 1.26) free manipulation. In addition, participants reported more efficiency (V = 434, p < .01) for automatic manipulation with large objects (Mdn = 1.88, M = 1.91, SD = .90) than small objects (Mdn = 1.25, M = 1.23, SD = 1.14). Moreover, they reported more efficiency in manipulating large objects by automatic manipulation than by free one both large (V = 520.00, p < .001; Mdn = .75, M = .92, DS = 1.19) and small (V = 610, p < .001; Mdn = .75, M = .89, DS = 1.08) objects. Although the stimulation dimension was significant in Wilcoxon's test for interaction (object manipulation x object size), no significance emerged from post hoc tests with B.H. corrections. Finally, no post hoc tests were performed for dependability and novelty dimensions.

As for excitement, there are no effects of the two variables and their interaction, as shown in the table below (Table 4.4).

Excitement quality scale				
Dimension	Analysis	V	z	p (B.H.)
	Object Size	361.50	.76	.497
Excitement	Object Manipulation	306.00	-0.69	.497
	Object Size x Object Manipulation	449.50	2.20	.084

*Table 4.4: Results of Wilcoxon tests on excitement qulity scale; (*p(B.H.) < .05).*

Positive scores were given in all experimental sessions. In fact, the results of the Wilcoxon test in relation to the median of the scale (Mdn = 4) showed high levels of excitement during the purchase of large items both by freely exploring them (V = 689.00, p < .001; Mdn = 5.40, M = 5.45, SD = 0.99) and through the anchor system (V = 812.00, p < .001; Mdn = 5.80, M = 5.59, SD = .92). The same results were obtained during the manipulation of small items with both automatic (V = 740.50, p < .001; Mdn = 5.20, M = 5.28, SD = 1.03) and free (V = 797.50, p < .001; Mdn = 5.60, M = 5.60, SD = .96) manipulation modalities.

Users' opinions on 3D Virtual Object

The table below shows the results of the Wilcoxon tests that emerged for each dimension of the presence in AR environment questionnaire (Table 4.5).

Presence in AR environment questionnaire				
Dimensions	Indipendent variable	V	z	p (B.H.)
Realness*	Object Size	312.00	-0.05	.97
	Object Manipulation	255.50	.47	.97
	Object Size x Object Manipulation	439.00	2.83	<.01
Spatial presence	Object Size	201.00	-1.42	.45
	Object Manipulation	302.00	-0.49	.83
	Object Size x Object Manipulation	293.00	.22	.83
Perceptual stress	Object Size	268.00	.39	.70
	Object Manipulation	248.50	1.04	.45
	Object Size x Object Manipulation	171.00	-1.26	.45

*Table 4.5: Results of Wilcoxon tests on user experience questionnaire; (*p(B.H.) < .05).*

Although an effect of interaction on the realness dimension emerged, post-hoc tests showed no statistically significant differences.

Wilcoxon tests in comparison with the median of the scale (Mdn = 4) performed, showed that people declared neither agreement nor disagreement to perceive virtual objects as physical objects during automatic manipulation of the small ones (V = 405.50, p = 28; Mdn = 4.50, M = 4.27, SD = 1.54), and during the free manipulation of the large objects (V = 313.0, p = .57; Mdn = 4.67, M = 4.18, SD = 1.38) and small ones (V = 508.00, p = .06; Mdn = 5.00, M = 4.51, SD = 1.46). Otherwise, when handling large objects in automatic mode, the products are perceived as physical (V = 559.50, p < .01; Mdn = 4.67, M = 4.58, SD = 1.25).

Otherwise, positive scores are given to spatial presence both during automatic manipulation of large objects (V = 759.00, p < .001; Mdn = 5.50, M = 5.44, SD = 1.27) and small objects (V = 639.50, p < .001; Mdn = 5.75, M = 5.60, SD = 1.27) and during free manipulation of large objects (V = 731.00, p < .001; Mdn = 5.50, M = 5.48, SD = 1.09) and small objects (V = 789.00, p < .001; Mdn = 6.00, M = 5.70, SD = 1.04). So, in all conditions, holograms at the perceptual level are placed in space and have their tridimensionality.

In addition, users report no perceptual stress during automatic manipulation of large objects (V = 89.00, p < .001, Mdn = 3.00, M = 3.05, SD = 1.84) and small objects (61.00, p < .001, Mdn = 3.25, M = 2.99, SD = 1.23), and during free manipulation of large objects (V = 27.50, p < .001; Mdn = 2.75, M = 2.89, SD = .98) and small objects (V = 31.50, p < .001; Mdn = 2.50, M = 2.70, SD = 1.05).

Moreover, regarding the quality of 3D virtual models, a series of Mann-Withey tests highlight that no differences emerged in terms of graphic quality (W = 2973, p = .44) between the two different typologies of products, namely interior furnishing (M = 6.51, SD = 1.65) and household appliances (Mdn = 6.59, M = 6.68, SD = 1.71). The same results were also obtained for the vividness (W = 3189, p = .97) of interior furnishing (Mdn = 6.67, M = 6.28, SD = 1.91) and household appliances (Mdn = 6.59, M = 6.30, SD = 2.00). Finally, no difference emerged (W = 3490.50, p = .32) in the graphic quality and vividness (W = 3500.50, p = .31) of large products (graphic quality: Mdn = 6.83, M = 6.73, SD = 1.63; vividness: (Mdn = 7.00, M = 6.14, SD = 2.01) and small ones (graphic quality: Mdn = 6.41, M= 6.47, SD = 1.73; vividness: Mdn = 6.33, M = 6.46, SD = 1.90).

Customer satisfaction

The table below shows the results of the Wilcoxon tests that emerged for each dimension of the customer satisfaction questionnaire (Table 4.6).

Customer Satisfaction Questionnaire				
Dimensions	Analysis	V	z	p (B.H.)
Overall Satisfaction	Object Size	122.50	1.11	.25
	Object Manipulation	168.50	1.84	.09
	Object Size x Object Manipulation	181.00	2.28	.06
Desire to stay	Object Size	193.00	1.67	.12
	Object Manipulation	204.00	1.12	.26
	Object Size x Object Manipulation	193.50	1.69	.12
Repatronage*	Object Size	199.00	- .09	.94
	Object Manipulation	256.00	- .15	.89
	Object Size x Object Manipulation	418.00	3.33	< .01
Luxury	Object Size	271.00	1.97	.15
	Object Manipulation	177.00	- .29	.78
	Object Size x Object Manipulation	146.00	-1.03	.45
Quality concern	Object Size	113.00	- .44	.67
	Object Manipulation	121.50	- .50	.67
	Object Size x Object Manipulation	87.50	-1.54	.38

*Table 4.6: Results of Wilcoxon tests on customer satisfaction questionnaire; (*p(B.H.) < .05).*

Post hoc results showed a higher level of repatronage (V = 87.00, p = .04) for free manipulation of small objects (Mdn = 5.75, M = 5.44, SD = 1.13) than for large objects (Mdn = 5.00, M = 5.09, SD = 1.15). No other statistically significant differences emerge, but the scores in the various conditions are all significantly higher than the median of the scale (Mdn = 4.00). Although an effect of interaction on the repatronage dimension emerged, post-hoc tests showed no statistically significant differences.

Furthermore, from Wilcoxon tests in comparison with the median of the scale (Mdn = 3), users reported being generally satisfied in all 4 conditions: automatic manipulation of large objects (V = 630, p < .001; Mdn = 4.00, M = 4.12, SD = .60), automatic manipulation of small objects (V = 604.50, p ; Mdn = 4.00, M = 3.78, SD = 89), free manipulation of large objects V = 604.50, p ; Mdn = 4.00, M = 4.12, SD = .60), and free manipulation of small objects V = 606.00, p ; Mdn = 4.00, M = 4.01, SD = .67). The same was observed for the desire to stay in the AR shop for all experimental conditions: automatic manipulation of large objects (V = 652.00, p ; Mdn =4.00, M = 4.01, SD = .67), automatic manipulation of small objects (V = 570.50, p ; Mdn = 4.00, M = 3.78, SD = .70), free manipulation of large objects (V = 466.50, p ; Mdn = 4.00, M = 3.75, SD = .90), and free manipulation of small objects (V = 536.00, p ; Mdn = 4.00, M = 3.83, SD = .84).

However, it emerges that the products are not considered luxury products by users: automatic manipulation of large objects (V = 27.50, p < .01; Mdn = 3.23, M = 3.26, SD = .51), automatic manipulation of small objects (V = 25.50, p < .049; Mdn = 3.20, M = 3.20, SD = .66), free manipulation of large objects GL (V = 40.50, p < .01; Mdn = 3.29, M = 3.29, SD = .57) and free manipulation of small objects (V = 12.00, p < .10; Mdn = 3.13, M = 3.13, SD = .65). Finally, participants neither agreed nor disagreed with respect to concern about the quality of physical products purchased in all the conditions (Mdn = 5): automatic manipulation of large objects (V = 189.50, p = .72; Mdn = 4.74, M = 4.97, SD = .84), automatic manipulation of small objects (V = 266.00, p = .72; Mdn = 4.75, M =5.12, SD =.85), free manipulation of large objects (V = 296.50, p = .72; Mdn = 5.00, M = 5.11, SD = .88), and free manipulation of small objects (V = 284.50, p = .72; Mdn = 4.40, M = 5.03, SD = .88).

4.3.2.1 Behavioral data

Purchasing success

Fisher's exact test was conducted to examine the association between experimental conditions and the correct purchase of products. The results showed no significant association, (p = 1.00), indicating that experimental conditions didn't have a significant impact on the likelihood of choosing the correct product.

Specifically, as shown in Table 4.7, users purchased the correct products in all the different augmented reality shops.

Purchase of products		
Object Size - Object Manipulation	correct (%)	error (%)
Large objects - Automatic scaling	100.00 %	0.00 %
Large objects - Free scaling	99.17 %	0.83 %
Small objects - Automatic scaling	98.33 %	1.67 %
Small objects - Free scaling	100.00 %	0.00 %

Table 4.7: Percentage of correct purchase of products.

Time invested in purchasing

The table below shows the time users spent shopping in the different experimental sessions. Comparisons were made to purchase times to see if there was an effect of object size and manipulation on time. A series of Shapiro-Wilk tests highlighted the non-normal distribution of the data. Therefore, non-parametric analyses were performed. Wilcoxon's tests (B.H. corrections) showed that there was no statistically significant difference in purchase time considering both object size ($V = 342.00$, $z = -0.91$, $p = .36$) and manipulation ($V = 258.00$, $z = -2.04$, $p = .12$) and the interaction between the two variables ($V = 313.00$, $z = -1.30$, $p = .30$). Thus, it turns out that users spent the same amount of time exploring the objects and making the purchase in the different sessions (Figure 4.3).

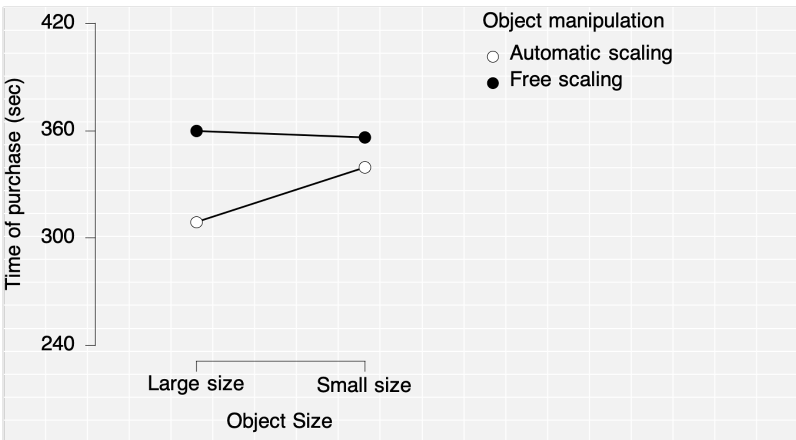

Figure 4.3: Time of purchase under different conditions.

Number and duration of manipulations

The analysis investigated whether there was an effect in the numbers and durations of manipulations.

The assumption of normality was verified using a series of Shapiro-Wilk tests. The significant results indicated a deviation from normality, therefore non-parametric analyses were conducted (Table 4.8).

Number and duration of manipulations				
Dependent variable	Independent variable	V	z	p (B.H.)
Number of manipulations*	Object Size	666.00	3.85	< .001
	Object Manipulation	45.00	4.91	< .001
	Object Size x Object Manipulation	207.00	2.32	.02
Duration of manipulations*	Object Size	657.00	3.32	< .001
	Object Manipulation	51.00	4.83	< .001
	Object Size x Object Manipulation	805.00	5.31	< .001

*Table 4.8: Results of Wilcoxon tests on number and duration of manipulations; (*p(B.H.) < .05).*

A series of post-hoc Wilcoxon test were performed. It was found that the number of manipulations performed is higher when the manipulation is free. Particularly, during free manipulation (V = 648.50, z = 3.60, p < .001) the participants manipulated the larger objects (Mdn = 7.08, M = 9.22, SD = 5.45) more than the smaller ones (Mdn = 6.58, M = 6.69, SD = 2.47).

Additionally, regarding the time spent to manipulate products the results are similar:

Users manipulated objects for longer when they had free manipulation. In addition, they spent more time in manipulating the larger objects (Mdn = 18.98, M = 25.71, SD = 17.46) compared to the smaller objects (Mdn = 18.73, M = 18.97, SD = 6.64); (Table 4.9).

Number and duration of manipulation: descriptive statistics						
Object Size - Object Manipulation	Number of manipulations			Duration of manipulations		
	Mnd	M	SD	Mdn	M	SD
Large objects - Automatic scaling	3.42	4.31	3.14	9.27	11.10	6.83
Large objects - Free scaling	7.08	9.22	5.45	18.98	25.71	17.46
Small objects - Automatic scaling	3.08	3.30	1.12	9.31	10.10	3.13
Small objects - Free scaling	6.58	6.69	2.47	18.73	18.97	6.42

Table 4.9: Descriptive statistics of numbers and durations of manipulations (median, mean and standard deviation).

Non-parametric correlations were applied to examine the relationship between the number and duration of manipulations during free manipulation of the objects. Spearman's rho indicated a significant positive correlation between time and the number of manipulations (r = .80, p < .001),

suggesting that as the number of manipulations increased, time spent on manipulation increased. Particularly, in particular, this correlation occurs during both free manipulations of large ($r = .61$, $p < .001$) and small objects ($r = .60$, $p < .001$).
Finally, correlations between the Need for touch (NFT scale, self-report data) and, respectively, numbers and duration of free object manipulations were examined using Spearman's rank correlation coefficient. The correlation was not statistically significant for both NFT and numbers of manipulations ($r = .09$, $p = .59$) and NFT and duration of manipulations ($r = -.09$, $p = .59$).

Overall user preferences

A series of Shapiro-Wilk tests highlighted the non-normal distribution of the data. Therefore, non-parametric analyses through Wilcoxon test (B.H. corrections) were performed. Users perceive it easier to manipulate a large object by means of automatic manipulation ($V = 725.50$, $p < .001$; $Mdn = 7.00$, $M = 5.98$, $SD = 1.64$), while for small objects they perceive using the free manipulation mode as easier ($V = 158.00$, $p < .01$; $Mdn = 2.00$, $M = 2.85$, $SD = 2.10$). In addition, participants reported that it is more pleasant to manipulate a large object through automatic manipulation ($V = 596.00$, $p < .001$; $Mdn = 7.00$, $M = 5.53$, $SD = 2.04$), mind a small object through the free one ($V = 120.00$, $p < .01$; $Mdn = 2.00$, $M = 2.90$, $SD = 1.92$). In addition, greater effectiveness emerged for large objects using the automatic manipulation ($V = 691.50$, $p < .001$; $Mdn = 7.00$, $M = 5.85$, $SD = 1.79$) while for small objects the free one ($V = 183.50$, $p < .01$; $Mdn = 2.00$, $M = 2.83$, $M = 2.18$). Finally, users perceive it as more useful to use automatic manipulation for large objects ($V = 388.50$, $p < .01$; $Mdn = 5.00$, $M = 4.98$, $SD = 1.98$) and free manipulation for small objects ($V = 89.00$, $p < .01$; $Mdn = 3.00$, $M = 3.05$, $SD = 1.92$).

4.3.4. Discussion

The present work investigated modalities to interact with 3D virtual products of different sizes during an augmented-reality shopping experience focusing on user experience and preferences. Specifically, in one case, virtual 3D objects of large and small sizes could be explored through hand gestures in their size true to the physical product (1:1 scale). On the other, the potential of augmented reality to enlarge or shrink products at will, without respecting the scale of the physical world, was exploited. To do this, participants were presented with a series of purchasing scenarios and asked to explore the products and purchase ones that satisfied the scenarios' requirements. Specifically, the scenarios involved exploring large (e.g., oven, armchair) or small (e.g., coffee maker, lamp) household appliances and home goods. Thus, participants explored large objects with automatic manipulation, large objects with free manipulation, small objects with automatic manipulation, and small objects with free manipulation.

During the AR experience, participants successfully purchased products under all conditions. In fact, users made the purchases demanded by the scenarios with high accuracy (98% to 100%) with no difference in their performance. In addition, people were satisfied with the augmented reality shopping experience and considered excited to explore the 3D virtual products of both large and small sizes in the modalities provided. These results highlighted how users reported an overall positive experience while manipulating the virtual products. However, significant differences emerged between different object sizes and manipulation modes in regard to user experience.

Consistently with our first hypothesis, manipulating large objects with the automatic resize by anchoring is considered more attractive than in the free manipulation modality. Also, the automatic modality was more attractive for large objects even compared with small ones. In addition, in terms of efficiency, manipulating large objects in the automatic modality was easier and required less effort as compared to free modality and small object manipulations (in both automatic and free modalities). Furthermore, the automatic modality was found to be more perspicuous as compared to the free one. In fact, it was easier for participants to learn how to manipulate virtual objects in the automatic resize by anchoring manipulation regardless of the size of the objects, as compared with the free manipulation modality both with large and small objects. Thus, manipulating large objects in automatic modality showed better usability; moreover, considering object size, automatic manipulation is considered more perspicuous with large products.

Behavioral data showed that people manipulate products more when it is possible to zoom them in and out according to their preferences. Particularly, they manipulate larger objects with more gestures and for a longer time than smaller ones. It could therefore be hypothesized that more interaction with the large product is required to find all necessary information and view the details of products. Nonetheless, the time spent to complete product purchases did not differ significantly among all the conditions, which means that the time to purchase products is not affected by the size of objects or manipulation modalities. This suggests that although participants during the automatic manipulation condition find it less challenging and easier to manipulate large products and find the information needed to purchase them, the exploration of large products does not turn out to be superficial or less careful.

In addition, after users explored and purchased large and small objects with both manipulation modalities, they reported their preferences. For large products users considered more effective, easy, useful, and pleasant the automatic manipulation modality, in line with our first hypothesis. Differently, for small products users considered more effective, easy, useful, and pleasant the free manipulation modality, in line with our second hypothesis.

Our results suggest that providing different modalities of manipulation depending on the size of the 3D products to be purchased allows for enriching the user experience, making it more satisfying and environmentally friendly. Because our gestures adapt to the affordances of objects (Pham et al., 2018) and, in our case, their sizes, it is appropriate to replicate what happens in the physical world in augmented reality environments as well.
In addition, users showed both a good level of acceptance regarding head-mounted displays such as HoloLens in shopping contexts and a desire to spend time in stores with AR applications.
Finally, users were neutral with respect to the level of realism of holograms; however, holograms are considered perceptually in physical space and their three-dimensionality was consistent, with no perceptual stress emerging. These results were consistent across all conditions, indicating that the size of the objects and the mode of manipulation did not affect participants' perception of the AR environment.

4.3.5. Guidelines for AR purchase applications of household products.

- Large 3D virtual products need a dimensional representation that is faithful to that of the physical product (1:1 scale).
- Exploration of large 3D virtual products should include positioning the object on the ground to understand its footprint.
- Exploration of small 3D virtual products requires the possibility to rotate the objects.
- Exploration of small 3D virtual products should include the ability to resize the object according to one's preferences.
- Exploration of products in settings such as stores and trade show displays needs a dedicated space free of clutter.

4.3.6 Conclusion

This work extends knowledge about interactions with 3D objects (Pham et al., 2018; Ortega et al., 2017) during a shopping experience with augmented reality. In addition, the findings helped to inform the design of AR applications in the retail sector, with particular reference to 3D product manipulation techniques.
We acknowledge the following limitation. The data collection has been constrained to the retail sale of household products (i.e., electrical appliances and furniture items). In the future, it is important to study whether products involving other sensory aspects (i.e., smell, taste) need different types of manipulations to enhance the shopping experience.

Finally, the considered sample of users was young adults, future studies should also evaluate aspects of object manipulation with people of different ages (e.g., elderly).

5. General conclusions

As emerged in previous chapters, the COVID-19 pandemic has brought to light the crucial role that new technologies play in our society in responding to difficulties and aiding in various sectors, such as entertainment, retail, education, and training. Despite the advancements made in these areas, there are still gaps and shortcomings when it comes to improving the user experience of these new technologies in different application areas.

The main goal of the present Ph.D. thesis was to understand how new interactive technologies can help in reducing the weaknesses that Covid-19 has brought out in different user scenarios, adopting a mixed approach in the Human Computer Interaction field. From a practical point of view, the three studies investigated in this project contribute to informing the design of interactive systems, offered by new technologies. In addition, it brings some guidelines and suggestions for enhancing experiences in the different areas examined.

The first study, Training in immersive VR, focused on the research of immersive virtual reality training, with a particular interest in flood emergencies that is a rising phenomenon. The goal was implementing an engaging and efficient immersive training to cope with floods through a human-centric design approach. This approach allowed us to individuate the simulation's design elements (e.g., realistic surroundings, multisensory feedback) and learning contents (e.g., procedural-practical knowledge). Specifically, in the early stage of co-design activities, two affinity diagram sessions and brainstorming were conducted with experts who had different theoretical backgrounds. The aim was to improve the quality of life of citizens exposed to the risk of floods. From these activities, a series of results emerged. First, the adopted methods helped in design a VR application focused on two main aspects of a flood situation: identifying the signs of a probable breakdown of the embankment, and adopting the correct behaviors during the emergency. Second, the key points that emerged allowed us to highlight a set of guidelines to support the design of VR simulation, namely the scenarios and system design, physical indicators of danger, and facing the emergency. In the second part of this work, the guidelines previously identified have been validated through an evaluation of the immersive VR experience by engaging target users. Furthermore, it highlighted the importance of providing a guided VR tutorial at the beginning of the experience. The tutorial allows citizens to familiarize themselves with the commands to be used in VR and with the environment in which they are immersed, fostering the experience. Overall, the effectiveness of the simulation in training users for the emergency, as well as the user engagement and relative interest in conducting such training with immersive VR rather than traditional methods, shows the potential of immersive environments. Thus, immersive VR offers the opportunity to conduct effective courses remotely, permitting the acquisition of useful skills to cope with emergencies when face-to-face courses are not possible. In addition, this technology could be implemented during in-person courses as a practical exercise.

The second study of the project expanded previous research about sharing remote experience between people in the entertainment field. The study investigated the effect of different modalities of computer-mediated communication on group dynamics during a social TV experience of watching an interactive digital narrative. Users who have experienced IDN as a group reported the intention to further explore similar experiences in the future. Indeed, our results are in line with the literature: computer-mediated interaction systems are related to positive phenomena also during IDN experience (e.g., status equalization effects, absence of flaming, high degree of enjoyment, and social conformity; Kelly, 2004; Laporte, 2010). Our finding suggested that while watching synchronous IDN in online and synchronous groups, audio and video-audio chats support a more frequent and better interaction between users than text chats, especially at the narrative crossroads. In addition, these CMC systems were more appreciated by the audience. Besides, it should be noted that the users who communicated using the audio channel achieved the narrative goal better than the other groups and single-players, completing the episode by selecting the ideal path in most cases. The findings of this study not only broaden the knowledge related to computer-mediated interaction during IDN watching, but also provide design suggestions for current streaming services to improve the performance and enjoyment of an IDN group experience. It is known that social TV provides an opportunity to build viewer engagement with programs, movies, and their advertisers (Nagy & Midha, 2014; Pynta et al., 2014). However, this study clearly showed the importance to exploite an auditory communication system rather than a visual one, during a share communication. Nowadays this service can be easily implemented on streaming platforms (that are currently characterized only by text chat) and in future applications thanks to the microphone integrated into the users' devices or the recent remote controls for TV. It is also essential to create contents that lead to a higher capacity of identification with the main character. The IDN design need to take into account a more comprehensive system of narrative choices and crossroads to increase the sense of autonomy of the users in the narrative decision-making (Roth, 2015).

The third study focused on the retail sector and the contribution of new technologies in this application area. The first data collection presented aimed to investigate the factors that facilitate or hinder the adoption of e-commerce by small businesses, as well as the potential integration of 3D web technology into their operations. Furthermore, we explored how operators in the sector of trade fairs consider the application of AR technology, which has yet to be investigated previously. As a result of the interviews, the obstacles identified in previous research about e-commerce were confirmed, mainly organizational and technological factors (Hong & Zhu, 2006;

Wymer & Regan, 2005). Additionally, the data collection translated the findings into practical guidelines to assist in designing user-friendly interfaces to enable small businesses to initiate and manage their e-commerce activities. These guidelines highlight the importance of easy and efficient information input and management to ensure seamless integration with all other work activities. Regarding the exhibition sector, professionals were generally receptive to integrating AR technology as it would allow them to exhibit a greater number of products at events reducing organization and transportation costs. Additionally, AR technology presents novel opportunities to explore products from different perspectives. However, it is recommended that AR technology should be employed as a complement to traditional stands, rather than a substitute for them, to enhance the value of trade fairs. Finally, the second data collection presented in this study expanded previous research about the customer experience with AR applications. In particular, the study investigated modalities to interact with 3D virtual products of different sizes during an augmented-reality shopping experience. The findings helped to inform the design of AR applications in the retail sector, with particular reference to 3D product manipulation techniques.

The findings of this Ph.D. project have emphasized the essential role that target user involvement plays in designing and evaluating various systems in different contexts of use. Moreover, users actively underlined the importance that new technologies have in various context of use to cope the fragility that COVID-19 has brought out in the physical environment.

In addition, the results obtained represent a first attempt facilitating remote activities that promote learning, sharing, entertainment, and shopping, unhindered by temporal or spatial constraints. The unprecedented circumstances brought about by the COVID-19 pandemic have brought to the forefront the potential of emerging technologies and the transformation of our activities through hybrid spaces, reshaping our perception of physical boundaries. Finally, the outcomes obtained contribute to providing useful indications that go beyond the emergency moment, helping target users with different individual characteristics or needs.

Despite the contribution offered by this project, it is important to acknowledge certain limitations. One primary limitation is associated with the predominantly homogeneous characteristics of the target users (e.g., age, nationality) involved in data collection. To address this limitation, future studies should focus on target users with various sociodemographic characteristics. Doing so enhances the external validity of these findings and contributes to a more inclusive and comprehensive body of knowledge regarding the impact of sociodemographic factors in the field of new technologies. Furthermore, another limitation of this project is that studies focused on evaluating specific experiences that are inherently linked to the content being examined. Consequently, the interpretation of the results should be contextualized and understood within the specific framework of the content. Therefore, caution should be exercised when generalizing the findings to other contexts or content domains. Future challenges will contribute to expanding these studies incorporating genuine situations and other realistic scenarios in order to offer other optimal solutions for the extended reality.

Reference

Albus, P., Vogt, A., & Seufert, T. (2021). Signaling in virtual reality influences learning outcome and cognitive load. *Computers & Education*, 166, 104154. https://doi.org/10.1016/j.compedu.2021.104154

Andersen, E., O'Rourke, E., Liu, Y. E., Snider, R., Lowdermilk, J., Truong, D., ... & Popovic, Z. (2012, May). The impact of tutorials on games of varying complexity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 59-68). https://doi.org/10.1145/2207676.2207687

Andersen, I. N. S. K., Kraus, A. A., Ritz, C., & Bredie, W. L. (2019). Desires for beverages and liking of skin care product odors in imaginative and immersive virtual reality beach contexts. *Food Research International*, *117*, 10-18. https://doi.org/10.1016/j.foodres.2018.01.027

Andersen, M. S., & Makransky, G. (2021). The validation and further development of a multidimensional cognitive load scale for virtual environments. *Journal of Computer Assisted Learning*, 37(1), 183-196. https://doi.org/10.1111/jcal.12478

Andersen, S. M., & Thorpe, J. S. (2009). An if–then theory of personality: Significant others and the relational self. *Journal of Research in Personality*, *43*(2), 163–170. https://doi.org/10.1016/j.jrp.2008.12.040

Anthony, C.M., Falzon, L., 2020. Understanding motivations to use online streaming services: integrating the technology acceptance model (TAM) and the uses and gratifications theory (UGT). Spanish J. Mark. ESIC. https://doi.org/10.1108/SJME-04-2020-0074

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, 55, 469– 480. https://doi.org/10.1037/0003-066X.55.5.469

Asadzadeh, A., Samad-Soltani, T., & Rezaei-Hachesu, P. (2021). Applications of virtual and augmented reality in infectious disease epidemics with a focus on the COVID-19 outbreak. *Informatics in medicine unlocked*, *24*, 100579. https://doi.org/10.1016/j.imu.2021.100579

Baboo, S., Kanna, Y., & Bennett, C. N. (2022). A Systematic Review on the Neuro-Cognitive Correlates of Game-Based Learning in Higher Education Learning Environments. *Handbook of Research on Acquiring 21st Century Literacy Skills Through Game-Based Learning*, 58-77. https://doi.org/10.4018/978-1-7998-7271-9.ch004

Baillie, L., Frohlich, P., & Schatz, R. (2007, June). Exploring social TV. In *2007 29th International Conference on Information Technology Interfaces* (pp. 215-220). IEEE.

Baltes, B. B., Dickson, M. W., Sherman, M. P., Bauer, C. C., & LaGanke, J. S. (2002). *Computer-mediated communication and group decision making: A meta-analysis.* https://doi.org/10.1006/obhd.2001.2961

Bamodu, O., & Ye, X. M. (2013). Virtual reality and virtual reality system components. *Advanced Materials Research*, 765–767, 1169–1172. https://doi.org/10.4028/www.scientific.net/AMR.765-767.1169

Baños, R. M., Botella, C., Garcia-Palacios, A., Villa, H., Perpiñá, C., & Alcaniz, M. (2000). Presence and reality judgment in virtual environments: a unitary construct?. *CyberPsychology & Behavior*, *3*(3), 327-335. https://doi.org/10.1089/10949310050078760

Barbosa Escobar, F., Petit, O., & Velasco, C. (2021). Virtual terroir and the premium coffee experience. *Frontiers in Psychology, 12,* 586983. https://doi.org/10.3389/fpsyg.2021.586983

Barhorst, J., McLean, G., Krey, N., Evanschitzky, H., & Javornik, A. (2021, June). Augmented reality experiences: exploring psychological, cognitive, and sensory aspects. In *2021 Academy of Marketing Science Annual Conference: Celebrating the Past and Future of Marketing*. https://doi.org/10.1007/978-3-030-95346-1_138

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Benvegnù, G., Furlan, M., Spagnolli, A., and Gamberini, L. (2019). "Emergency fire training in virtual environments: effect of immersion and interaction," in *Oral Presentation at 19th Congress of European Association of Work and Organizational Psychology (EAWOP)* (Turin).

Berg, E. C., Trost, M., Schneider, I. E., & Allison, M. T. (2001). Dyadic exploration of the relationship of leisure satisfaction, leisure time, and gender to relationship satisfaction. *Leisure sciences*, *23*(1), 35-46. https://doi.org/10.1080/01490400150502234

Bettelli, A., Orso, V., Pluchino, P., and Gamberini, L. (2019). "An enriched visit to the botanical garden: co-designing tools and contents," in *Proceedings of the 13th Biannual Conference of the Italian SIGCHI Chapter: Designing the Next Interaction (Padova)*. https://doi.org/10.1145/3351995.3352034

Bevan, N., Carter, J., Earthy, J., Geis, T., & Harker, S. (2018, July). What are user requirements? Developing an ISO standard. In *International Conference on Human-Computer Interaction* (pp. 3-13). Springer, Cham. https://doi.org/10.1007/978-3-319-91238-7_1

Billewar, S. R., Jadhav, K., Sriram, V. P., Arun, A., Abdul, S. M., Gulati, K., & Bhasin, N. K. K. (2021). The rise of 3D E-Commerce: the online shopping gets real with virtual reality and augmented reality during COVID-19. *World Journal of Engineering*. https:///doi.org/10.1108/WJE-06-2021-0338

Blackler, A., Zelenko, O., Chamorro-Koc, M., Rittenbruch, M., & Sade, G. (2018). Design research: methodological innovation through messiness. In New Directions in Third Wave Human-Computer Interaction: Volume 2-Methodologies (pp. 131-154). Springer, Cham. https://doi.org/10.1007/978-3-319-73374-6_8

Boehm, B. W., Brown, J. R., & Lipow, M. (1976, October). Quantitative evaluation of software quality. In *Proceedings of the 2nd international conference on Software engineering* (pp. 592-605).

Boger, J., Eisapour, M., Cao, S., and Domenicucci, L. (2018). "Participatory design of a virtual reality exercise for people with mild cognitive impairment," in *Conference on Human Factors in Computing Systems - Proceedings, 2018-April* (Montreal, QC), 1–9. https://doi.org/10.1145/3170427.3174362

Botella, C., Bretón-López, J., Quero, S., Baños, R. M., and García-Palacios, A. (2010). Treating cockroach phobia with augmented reality. *Behav. Ther.* 41, 401–413. https://doi.org/10.1016/j.beth.2009.07.002

Bozgeyikli, E., Raij, A., Katkoori, S., & Dubey, R. (2016, October). Point & teleport locomotion technique for virtual reality. In *Proceedings of the 2016 annual symposium on computer-human interaction in play* (pp. 205-216). https://doi.org/10.1145/2967934.2968105

Bretón-López, J., Quero, S., Botella, C., García-Palacios, A., Baños, R. M., and Alcañiz, M. (2010). An augmented reality system validation for the treatment of cockroach phobia. *Cyberpsychol. Behav. Soc. Netw.* 13, 705–710. https://doi.org/10.1089/cyber.2009.0170

Buttussi, F., & Chittaro, L. (2017). Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE transactions on visualization and computer graphics*, *24*(2), 1063-1076. https://doi.org/10.1109/TVCG.2017.2653117

Çakiroğlu, Ü., & Gökoğlu, S. (2019). Development of fire safety behavioral skills via virtual reality. *Computers and Education*, *133*(January), 56–68. https://doi.org/10.1016/j.compedu.2019.01.014

Çakıroğlu, Ü., & Gökoglu, S. (2019). Development of fire safety behavioral skills via virtual reality. *Comput. Educ*., 133, 56-68. https://doi.org/10.1016/j.compedu.2019.01.014

Carte, T. A., Chidambaram, L. e Becker, A. (2006). Emergent Leadership in Self-Managed Virtual Teams. *Group Decision and Negotiation*, 15, 323-343. https://doi.org/10.1007/s10726-006-9045-7

Chao, C. J., Wu, S. Y., Yau, Y. J., Feng, W. Y., & Tseng, F. Y. (2017). Effects of three-dimensional virtual reality and traditional training methods on mental workload and training performance. H*uman Factors and Ergonomics in Manufacturing & Service Industries*, 27(4), 187-196. https://doi.org/10.1002/hfm.20702

Checa, D., & Bustillo, A. (2020a). A review of immersive virtual reality serious games to enhance learning and training. *Multimedia Tools and Applications*, 79(9), 5501-5527. https://doi.org/10.1007/s11042-019-08348-9

Checa, D., & Bustillo, A. (2020b). Advantages and limits of virtual reality in learning processes: Briviesca in the fifteenth century. *Virtual Reality*, 24(1), 151-161. https://doi.org/10.1007/s10055-019-00389-7

Chen, Yunqiang & Wang, Qing & Chen, Hong & Song, Xiaoyu & Tang, Hui & Tian, Mengxiao. (2019). An overview of augmented reality technology. *Journal of Physics: Conference Series*. 1237. 022082. https://doi.org/10.1088/1742-6596/1237/2/022082

Chittaro, L., & Buttussi, F. (2015). Assessing Knowledge Retention of an Immersive Serious Game vs. a Traditional Education Method in Aviation Safety. *IEEE transactions on visualization and computer graphics*, 21(4), 529–538. https://doi.org/10.1109/TVCG.2015.2391853

Chittaro, L., & Buttussi, F. (2015). Assessing knowledge retention of an immersive serious game vs. A traditional education method in aviation safety. *IEEE Transactions on Visualization and Computer Graphics*, *21*(4), 529–538. https://doi.org/10.1109/TVCG.2015.2391853

Chittaro, L., Buttussi, F., & Zangrando, N. (2014, November). Desktop virtual reality for emergency preparedness: user evaluation of an aircraft ditching experience under different fear arousal conditions. In Proceedings of the 20th ACM symposium on virtual reality software and technology (pp. 141-150). https://doi.org/10.1145/2671015.2671025

Cipresso, P., Giglioli, I. A. C., Raya, M. A., & Riva, G. (2018). The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature. *Frontiers in psychology*, 2086. https://doi.org/10.3389/fpsyg.2018.02086

Cohen, E. L., & Lancaster, A. L. (2014). Individual differences in in-person and social media television coviewing: The role of emotional contagion, need to belong, and coviewing orientation. *Cyberpsychology, Behavior, and Social Networking*, *17*(8), 512-518. https://doi.org/10.1089/cyber.2013.0484

D'Amico, A., Bernardini, G., Lovreglio, R., & Quagliarini, E. (2022). A Non-Immersive Virtual Reality Serious Game Application for Flood Safety Training. Available at *SSRN 4110990*. https://doi.org/10.2139/ssrn.4110990

Dargan, S., Bansal, S., Kumar, M., Mittal, A., & Kumar, K. (2022). Augmented Reality: A *Comprehensive Review. Archives of Computational Methods in Engineering*, 1-24. https://doi.org/10.1007/s11831-022-09831-7

Deci, E.L. and Ryan, R.M. 2000. The "What" and "Why" of Goal Pursuits: Human Needs and the Self Determination of Behavior. Psychological Inquiry. 11, 4 (Oct. 2000), 227–268 https://doi.org/10.1207/S15327965PLI1104_01

Decreto del Presidente del Consiglio dei Ministri (2020). https://www.lavoro.gov.it/documenti-e-norme/normative/Documents/2020/DPCM-26-aprile-2020.pdf

Dezfuli, N., Khalilbeigi, M., Mühlhäuser, M., & Geerts, D. (2011, June). A study on interpersonal relationships for social interactive television. In *Proceedings of the 9th European Conference on Interactive TV and Video* (pp. 21-24). https://doi.org/10.1145/2000119.2000123

Dix, A., Finlay, J., Abowd, G. D., and Beale, R. (2004*). Human-Computer Interaction, 3rd Edn.* Upper Saddle River, NJ: Pearson Prentice Hall. ISBN: 978-0-13-046109-4

Doherty-Sneddon, G., Anderson, A., O'Malley, C., Langton, S., Garrod, S. e Bruce, V. (1997). Face-to-face and video-mediated communication: A comparison of dialogue structure and task performance. *Journal of Experimental Psychology: Applied*, 3, 105-125., 87(1), 156-179. https://doi.org/10.1037/1076-898X.3.2.105

Dubrovsky, V. J., Kiesler, S. e Sethna, B. N. (1991). The equalization phenomenon: Status effects in computer-mediated and face-to-face decision making groups. *Human Computer Interaction, 6,* 119-146. https://doi.org/10.1207/s15327051hci0602_2

Erdfelder, E., Faul, F., and Buchner, A. (1996). Gpower: A general power anal- ysis program. *Behavior research methods, instruments, & computers*, 28(1):1– 11.

Eisapour, M., Cao, S., and Boger, J. (2020). Participatory design and evaluation of virtual reality games to promote engagement in physical activity for people living with dementia. *J. Rehabil. Assist. Technol.* Eng. 7:205566832091377. https://doi.org/10.1177/2055668320913770

Englmeier, D., Dörner, J., Butz, A., & Höllerer, T. (2020, March). A tangible spherical proxy for object manipulation in augmented reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 221-229). IEEE. https://doi.org/10.1109/VR46266.2020.00041

Escobar F, Petit O and Velasco C (2021) Virtual Terroir and the Premium Coffee Experience. *Front. Psychol*. 12:586983. https://doi.org/10.3389/fpsyg.2021.586983

Farshid, M., Paschen, J., Eriksson, T., & Kietzmann, J. (2018). Go boldly!: Explore augmented reality (AR), virtual reality (VR), and mixed reality (MR) for business. *Business Horizons*, *61*(5), 657-663. https://doi.org/10.1016/j.bushor.2018.05.009

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

Feng, Z., González, V. A., Mutch, C., Amor, R., and Cabrera-Guerrero, G. (2021). Instructional mechanisms in immersive virtual reality serious games: earthquake emergency training for children. *J. Comput. Assist. Learn.* 37, 542–556. https://doi.org/10.1111/jcal.12507

Fernandes, B., Biswas, U. N., Mansukhani, R. T., Casarín, A. V., & Essau, C. A. (2020). The impact of COVID-19 lockdown on internet use and escapism in adolescents. *Revista de psicología clínica con niños y adolescentes*, *7*(3), 59-65. https://doi.org/10.21134/rpcna.2020.mon.2056

Fitzpatrick, T. R. (2009). The quality of dyadic relationships, leisure activities and health among older women. *Health Care for Women International*, *30*(12), 1073-1092. https://doi.org/10.1080/07399330903199367

Flobak, E., Wake, J. D., Vindenes, J., Kahlon, S., Nordgreen, T., and Guribye, F. (2019). “Participatory design of VR scenarios for exposure therapy,” in *Conference on Human Factors in Computing Systems – Proceedings* (Glasgow), 1–12. https://doi.org/10.1145/3290605.3300799

Fogg, B. J. (2002). *Persuasive technology: using computers to change what we think and do*. Ubiquity, 2002 (December), 2. https://doi.org/10.1145/764008.763957

Forlizzi, J., & Battarbee, K. (2004). Understanding experience in interactive systems. Paper presented at the *Proceedings of the 5th conference on designing interactive systems: Processes, practices, methods, and techniques*, Cambridge, MA https://doi.org/10.1145/1013115.1013152

Frauke Kuhn, Marcel Lichters, Nina Krey. (2020). The touchy issue of produce: Need for touch in online grocery retailing. Journal of Business Research 117 (2020) 244–255. https://doi.org/10.1016/j.jbusres.2020.05.017

Freina, L., & Ott, M. (2015, April). A literature review on immersive virtual reality in education: state of the art and perspectives. In *The international scientific conference elearning and* *software for education* (Vol. 1, No. 133, pp. 10-1007). https://doi.org/10.12753/2066-026X-21-020

Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*(11), 1325–1330. https://doi.org/10.1111/2041-210X.12584

Frommel, J., Fahlbusch, K., Brich, J., & Weber, M. (2017, October). 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). https://doi.org/10.1145/3116595.3116610

Fujimi, T., and Fujimura, K. (2020). Testing public interventions for flash flood evacuation through environmental and social cues: the merit of virtual reality experiments. *Int. J. Disaster Risk Reduc.* 50:101690. https://doi.org/10.1016/j.ijdrr.2020.101690

Gamberini, L., Bettelli, A., Benvegnù, G., Orso, V., Spagnolli, A., & Ferri, M. (2021). Designing "Safer Water." A Virtual Reality Tool for the Safety and the Psychological Well-Being of Citizens Exposed to the Risk of Natural Disasters. *Frontiers in psychology*, 12, 674171. https://doi.org/10.3389/fpsyg.2021.674171

Gamberini, L., Cottone, P., Spagnolli, A., Varotto, D., and Mantovani, G. (2003). Responding to a fire emergency in a virtual environment: different patterns of action for different situations. *Ergon*. 46, 842–858. https://doi.org/10.1080/0014013031000111266

Geerts, D. (2006, October). Comparing voice chat and text chat in a communication tool for interactive television. In *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles* (pp. 461-464). https://doi.org/10.1145/1182475.1182537

Gemmano, C. G., Manuti, A., & Giancaspro M. L. (2022). Does workplace training still work? Testing the efficacy of online and in-presence training modalities in the post-pandemic scenario [Abstract]. In Book of Abstract; Proceedings of the *30th Congress of the Italian Psychology Association*, (IPA 2022).

Gloumeau, P. C., Stuerzlinger, W., & Han, J. (2020). Pinnpivot: Object manipulation using pins in immersive virtual environments. *IEEE transactions on visualization and computer graphics*, *27*(4), 2488-2494. https://doi.org/ https://doi.org/10.1109/tvcg.2020.2987834

Guthrie, C., Fosso-Wamba, S., & Arnaud, J. B. (2021). Online consumer resilience during a pandemic: An exploratory study of e-commerce behavior before, during and after a COVID-19 lockdown. *Journal of Retailing and Consumer Services*, *61*, 102570. https://doi.org/10.1016/j.jretconser.2021.102570

Hagita, K., Kodama, Y., and Takada, M. (2020). Simplified virtual reality training system for radiation shielding and measurement in nuclear engineering. *Progr. in Nucl. Energy* 118:103127. https://doi.org/10.1016/j.pnucene.2019.103127

Halcomb, E. J., & Davidson, P. M. (2006). Is verbatim transcription of interview data always necessary?. *Applied nursing research*, *19*(1), 38-42. https://doi.org/10.1016/j.apnr.2005.06.001

Han, J., Zheng, Q., & Ding, Y. (2021). Lost in virtual reality? Cognitive load in high immersive VR environments. *JAIT*, 12(4). https://doi.org/10.12720/jait.12.4.302-310

Harboe, G., Massey, N., Metcalf, C., Wheatley, D., & Romano, G. (2008). The uses of social television. *Computers in entertainment (CIE)*, *6*(1), 1-15. https://doi.org/10.1145/1350843.1350851

Hartson, R., and Pyla, P. S. (2012). *The UX Book : Process and Guidelines for Ensuring a Quality User Experience*. Walthman, MA: Elsevier.

Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction*, 13(4), 481-499. http://dx.doi.org/10.1207/S15327590IJHC1304_07

He, Z., & Yang, X. (2014, November). Hand-based interaction for object manipulation with augmented reality glasses. In *Proceedings of the 13th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry* (pp. 227-230). http://dx.doi.org/10.1145/2670473.2670505

Hertel, J., Karaosmanoglu, S., Schmidt, S., Bräker, J., Semmann, M., & Steinicke, F. (2021, October). A taxonomy of interaction techniques for immersive augmented reality based on an iterative literature review. In *2021 IEEE international symposium on mixed and augmented reality (ISMAR)* (pp. 431-440). IEEE. https://doi.org/10.1109/ISMAR52148.2021.00060.

Hilken, T., Chylinski, M., Keeling, D. I., Heller, J., de Ruyter, K., & Mahr, D. (2022). How to strategically choose or combine augmented and virtual reality for improved online experiential retailing. *Psychology &Marketing,* *39*(3), 495–507. https://doi.org/10.1002/mar.21600

Hong, W., & Zhu, K. (2006). Migrating to internet-based e-commerce: Factors affecting e-commerce adoption and migration at the firm level. *Information & management*, *43*(2), 204-221. https://doi.org/10.1016/j.im.2005.06.003

Huang, M.-H. Designing Website Attributes to Induce Experiential Encounters. *Comput. Hum. Behav*. 2003, 19, 425–442. https://doi.org/10.1016/S0747-5632(02)00080-8

ISO - ISO 9241-210:2010 - Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. (2019). Iso.Org. https://www.iso.org/standard/52075.html

ISO. (2018). *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. Iso 9241-11:2018(E). https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en

Issa, T., & Isaias, P. (2015). Chapter 5: Sustainable Design. *Sustainable Design.* Springer London.

Jääskeläinen, R. (2010). Think-aloud protocol. *Handbook of translation studies*, *1*, 371-374.

Jílková, P. & Králová, P., 2021. Digital consumer behaviour and ecommerce trends during the COVID-19 crisis. *International Advances in Economic Research*, Vol. 27, No. 1, pp 83-85. https://doi.org/10.1007/s11294-021-09817-4

Juan, M. C., Alcaniz, M., Monserrat, C., Botella, C., Baños, R. M., and Guerrero, B. (2005). Using augmented reality to treat phobias. *IEEE Comput. Graph. Appl.* 25, 31–37. https://doi.org/10.1109/MCG.2005.143

Jussila, I., Tarkiainen, A., Sarstedt, M., & Hair, J.F. 2015. Individual psycho?logical ownership: concepts, evidence, and implications for research in marketing. *J. Mark. Theory Pract. 23* (2), 121–139. https://doi.org/10.1080/10696 679.2015.1002330

Kang, H. J., Shin, J. H., & Ponto, K. (2020, March). A comparative analysis of 3d user interaction: How to move virtual objects in mixed reality. In *2020 IEEE conference on virtual reality and 3D user interfaces (VR)* (pp. 275-284). IEEE. https://doi.org/10.1109/VR46266.2020.00047

Kangas, Jari & Kumar, Sriram & Mehtonen, Helena & Järnstedt, Jorma & Raisamo, Roope. (2022). Trade-Off between Task Accuracy, Task Completion Time and Naturalness for Direct Object Manipulation in Virtual Reality. *Multimodal Technologies and Interaction*. 6. 6. https://doi.org/10.3390/mti6010006

Kao, D., Magana, A. J., & Mousas, C. (2021). Evaluating Tutorial-Based Instructions for Controllers in Virtual Reality Games. *Proceedings of the ACM on Human-Computer Interaction*, 5(CHI PLAY), 1-28. https://doi.org/10.1145/3474661

Kardong-Edgren, S. (Suzie), Farra, S. L., Alinier, G., & Young, H. M. (2019). A Call to Unify Definitions of Virtual Reality. *Clinical Simulation in Nursing*, *31*, 28–34. https://doi.org/10.1016/j.ecns.2019.02.006

Ke, P., Keng, K. N., Jiang, S., Cai, S., Rong, Z., and Zhu, K. (2019). "Embodied weather: promoting public understanding of extreme weather through immersive multi-sensory virtual reality," in *Proceedings - VRCAI 2019: 17th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry* (Brisbane, QLD), 1–2. https://doi.org/10.1145/3359997.3365718

Kelly, J. R., & Loving, T. J. (2004). Time pressure and group performance: Exploring underlying processes in the attentional focus model. *Journal of experimental social psychology*, *40*(2), 185-198. https://doi.org/10.1016/S0022-1031(03)00094-5

Kent, S. L. (2001). The Ultimate History of Video Games. *Prima Communications*, Inc.,Rocklin, CA.

Kim, J., Merrill Jr, K., Collins, C., & Yang, H. (2021). Social TV viewing during the COVID-19 lockdown: The mediating role of social presence. *Technology in society*, *67*, 101733. https://doi.org/10.1016%2Fj.techsoc.2021.101733

Kim, J., Song, H., & Lee, S. (2018). Extrovert and lonely individuals' social TV viewing experiences: A mediating and moderating role of social presence. *Mass Communication and Society*, *21*(1), 50-70. https://doi.org/10.1080/15205436.2017.1350715

Kim, Y. M., Rhiu, I., & Yun, M. H. (2020). A systematic review of a virtual reality system from the perspective of user experience. *International Journal of Human–Computer Interaction*, 36(10), 893-910. https://doi.org/10.1080/10447318.2019.1699746

Kinateder, M., Ronchi, E., Nilsson, D., Kobes, M., Müller, M., Pauli, P., et al. (2014). "Virtual reality for fire evacuation research,," in *Proceedings of the 2014 Federated Conference on Computer Science and Information Systems*, Vol. 2 (Warsaw), 313–321. https://doi.org/10.15439/2014F94

Klimmt, C., Hartmann, T., and Frey, A. (2007). Effectance and control as determinants of video game enjoyment. *CyberPsychology & Behavior*, 10(6):845–7. http://dx.doi.org/10.1089/cpb.2007.9942

Ko, E., Costello, J. P., & Taylor, C. R. (2019). What is a luxury brand? A new definition and review of the literature. *Journal of Business Research, 99*, 405-413. https://doi.org/10.1016/j.jbusres.2017.08.023

Koenitz, H. (2015, November). Design approaches for interactive digital narrative. In *International Conference on Interactive Digital Storytelling* (pp. 50-57). Springer, Cham. https://doi.org/10.1007/978-3-319-27036-4_5

Koenitz, H., & Knoller, N. (2015). Interactive Digital Narratives for iTV and Online Video. *Handbook of Digital Games and Entertainment Technologies*, Springer, 1- 30. https://doi.org/10.1007/978-981-4560-52-8_44-1

Koenitz, H., & Louchart, S. (2015). Practicalities and Ideologies, (Re)-Considering the Interactive Digital Narrative Authoring Paradigm. In *FDG*.

Kortum, P. T., & Bangor, A. (2013). Usability ratings for everyday products measured with the System Usability Scale. I*nternational Journal of Human- Computer Interaction, 29*(2), 67-76. https://doi.org/10.1080/10447318.2012.681221

Kumar, V., Alshazly, H., Idris, S. A., & Bourouis, S. (2021). Evaluating the impact of covid-19 on society, environment, economy, and education. *Sustainability*, *13*(24), 13642. https://doi.org/10.3390/su132413642

Lai, J. S., Chang, W. Y., Chan, Y. C., Kang, S. C., and Tan, Y. C. (2011). Development of a 3D virtual environment for improving public participation: case study - the Yuansantze Flood Diversion Works Project. *Adv. Eng. Inform.* 25, 208–223. https://doi.org/10.1016/j.aei.2010.05.008

Laporte, L., van Nimwegen, C., & Uyttendaele, A. J. (2010, October). Do people say what they think: Social conformity behavior in varying degrees of online social presence. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries* (pp. 305-314). ACM. https://doi.org/10.1145/1868914.1868951

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In *HCI and Usability for Education and Work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, November 20-21, 2008. Proceedings 4* (pp. 63-76). Springer Berlin Heidelberg. http://dx.doi.org/10.1007/978-3-540-89350-9_6

LaValle, S. M. (2017) *Virtual Reality*. Cambridge: Cambridge University Press.

Leena Lakra P, Verma P (2017) Augmented reality in marketing: role and applications. 8(11):74–81. https://www.academia.edu/36282179/Augmented_Reality_in_Marketing_Role_and_Applications.

Leskens, J. G., Kehl, C., Tutenel, T., Kol, T., Haan, G., de Stelling, G., et al. (2017). An interactive simulation and visualization tool for flood analysis usable for practitioners. *Mitig. Adapt. Strateg. Glob. Change* 22, 307–324. https://doi.org/10.1007/s11027-015-9651-2

Li, C., Liang, W., Quigley, C., Zhao, Y., & Yu, L. (2017). Earthquake Safety Training through Virtual Drills. *IEEE Transactions on Visualization and Computer Graphics*, 23 (4), 1275-1284. https://doi.org/10.1109/TVCG.2017.2656958

Lin, J. S., Sung, Y., & Chen, K. J. (2016). Social television: Examining the antecedents and consequences of connected T.V. viewing. *Computers in Human Behavior*, *58*, 171-178. https://doi.org/10.1016/j.chb.2015.12.025

Linja-aho, M. (2006). Creating a framework for improving the learnability of a complex system. *Human Technology*, 2(2), 202-224. http://dx.doi.org/10.17011/ht/urn.2006519

Liu, X. J., Zhong, D. H., Tong, D. W., Zhou, Z. Y., Ao, X. F., and Li, W. Q. (2018). Dynamic visualisation of storm surge flood routing based on three-dimensional numerical simulation. *J. Flood Risk Manage.* 11, S729–S749. https://doi.org/10.1111/jfr3.12252

Löber, Andreas; Grimm, Sibylle; Schwabe, Gerhard (2006). *Audio vs chat: Can media speed explain the differences in productivity?* In: 14th European Conference on Information Systems, ECIS 2006, Göteborg, 12 June 2006 - 14 June 2006. https://doi.org/10.5167/uzh-61415

Loup, G., and Loup-Escande, E. (2019). Effects of Travel modes on performances and user comfort: a comparison between ArmSwinger and Teleporting. *Int. J. Hum. Comput. Interact.* 35, 1270–1278. https://doi.org/10.1080/10447318.2018.1519164

Lucero, A. (2015). “Using affinity diagrams to evaluate interactive prototypes,” *IFIP Conference on Human-Computer Interaction* (Bamberg), 231–248. https://dx.doi.org/10.1007/978-3-319-22668-2_19

Lukosch, H., Van Ruijven, T., and Verbraeck, A. (2012). “The participatory design of a simulation training game,” in *Proceedings - Winter Simulation Conference* (Harteveld), 1–11. https://doi.org/10.1109/WSC.2012.6465218

Lyk, P. B., Majgaard, G., Vallentin-Holbech, L., Guldager, J. D., Dietrich, T., Rundle-Thiele, S., et al. (2020). Co-designing and learning in virtual reality: development of tool for alcohol resistance training. *Electron. J. E-Learn.* 18, 213–228. https://doi.org/10.34190/EJEL.20.18.3.002

Macaranas, A., Venolia, G., Inkpen, K., & Tang, J. (2013). Sharing Experiences over Video: watching video programs together at a distance. In *Human-Computer Interaction–INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part IV 14* (pp. 73-90). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40498-6_5

Macchione, F., Costabile, P., Costanzo, C., and De Santis, R. (2019). Moving to 3-D flood hazard maps for enhancing risk communication. *Environ. Model. Softw.* 111, 510–522. https://doi.org/10.1016/j.envsoft.2018.11.005

Mahima, M., Shanthi Priya, R., Rajagopal, P. and Pradeepa, C. (2022), "Impact of Covid-19 on the built environment", Frontiers in Engineering and Built Environment, Vol. 2 No. 2, pp. 69-80. https://doi.org/10.1108/FEBE-09-2021-0040

Maital, S., & Barzani, E. (2020). The global economic impact of COVID-19: A summary of research. *Samuel Neaman Institute for National Policy Research*, *2020*, 1-12.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 60, 225– 236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Mason, V., Andrews, H., and Upton, D. (2010). The psychological impact of exposure to floods. *Psychol. Health Med.* 15, 61–73. https://doi.org/10.1080/13548500903483478

McIntosh, M. J., & Morse, J. M. (2015). Situating and constructing diversity in semi-structured interviews. *Global qualitative nursing research*, *2*. https://doi.org/10.1177/2333393615597674

Metcalf, C., Harboe, G., Tullio, J., Massey, N., Romano, G., Huang, E. M., & Bentley, F. (2008). Examining presence and lightweight messaging in a social television experience. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, *4*(4), 1-16. https://doi.org/10.1145/1412196.1412200

Michael, D. & Chen, S. *Serious Games: Games That Educate, Train, and Inform*; Thomson Course Technology: Boston, MA, USA, 2006.

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & education*, 56(3), 769-780. https://doi.org/10.1016/j.compedu.2010.10.020

Millen, D. R. (2000). Rapid ethnography: time deepening strategies for HCI field research. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 280-286). https://doi.org/10.1145/347642.347763

Mol, J. M., Botzen, W. W., & Blasch, J. E. (2022). After the virtual flood: Risk perceptions and flood preparedness after virtual reality risk communication. *Judgment and Decision Making*, 17(1), 189-214. http://dx.doi.org/10.1017/S1930297500009074

Monteiro, A. C. (2020). Interactive Digital Narratives: A Close Reading Of Bandersnatch To Analyze The Aesthetic Relations With The User. *Media Literacy and Academic Research*, 3(1), 57-72.

Morélot, S., Garrigou, A., Dedieu, J., & N'Kaoua, B. (2021). Virtual reality for fire safety training: Influence of immersion and sense of presence on conceptual and procedural acquisition. *Computers & Education*, 166, 104145. https://doi.org/10.1016/j.compedu.2021.104145

Morganti, F., & Riva, G. (2006). *Conoscenza, comunicazione e tecnologia: aspetti cognitivi della realtà virtuale.* LED Edizioni Universitarie.

Mulla, T. (2022). Assessing the factors influencing the adoption of over-the-top streaming platforms: A literature review from 2007 to 2021. *Telematics and Informatics*, 101797. https://doi.org/10.1016/j.tele.2022.101797

Müller, C. (2017). How recurrent gestures mean: Conventionalized contexts-of-use and embodied motivation. *Gesture*, *16*(2), 277-304. https://doi.org/10.1075/gest.16.2.05mul

Murray, J. H. (2012). Transcending Transmedia: Emerging Story Telling Structures for the Emerging Convergence Platforms (pp. 1–6). *Presented at the EuroITV 2012: 10th European Conference on Interactive Tv and Video*, New York: ACM. https://doi.org/10.1145/2325616.2325618

Nagy, J., & Midha, A. (2014). The value of earned audiences: How social interactions amplify T.V. impact: What programmers and advertisers can gain from earned social impressions. *Journal of Advertising Research*, *54*(4), 448-453. https://doi.org/10.2501/JAR-54-4-448-453

Nielsen, J. (1993). *Usability engineering*. San Francisco: Morgan Kaufmann.

Nielsen, J., and Norman, D. (2014) The Definition of User Experience, www.nngroup.com/ articles/definition-user-experience/ (accessed July 2, 2014).

Nofal, A., Alfayyad, I., Khan, A., Al Aseri, Z., & Abu-Shaheen, A. (2018). Knowledge, attitudes, and practices of emergency department staff towards disaster and emergency preparedness at tertiary health care hospital in central Saudi Arabia. *Saudi medical journal*, 39(11), 1123. https://doi.org/10.15537/smj.2018.11.23026

Nukarinen, T., Kangas, J., Rantala, J., Koskinen, O., and Raisamo, R. (2018). "Evaluating ray casting and two gaze-based pointing techniques for object selection in virtual reality," in *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST* (Tokyo), 17–18. https://doi.org/10.1145/3281505.3283382

Oliver, R. L., & Swan, J. E. (1989). Consumer perceptions of interpersonal equity and satisfaction in transactions: A field survey approach. *Journal of Marketing*, *53*(2), 21–35. https://doi.org/10.2307/1251411

Ortega, F. R., Galvan, A., Tarre, K., Barreto, A., Rishe, N., Bernal, J., ... & Thomas, J. L. (2017, March). Gesture elicitation for 3D travel via multi-touch and mid-Air systems for procedurally generated pseudo-universe. In *2017 IEEE symposium on 3D User Interfaces (3DUI)* (pp. 144-153). IEEE. https://doi.org/10.1109/3DUI.2017.7893331

Ozok, A. A., & Komlodi, A. (2007, October). Do looks really matter? The effect of 3-dimensional product representations on the customers' buying decision in electronic commerce. In *Proceedings of the Human Factors And Ergonomics Society Annual Meeting* (Vol. 51, No. 17, pp. 1050-1052). Sage CA: Los Angeles, CA: SAGE Publications.

Ozturkcan, S. (2021). Service innovation: Using augmented reality in the IKEA Place app. *Journal of Information Technology Teaching Cases*, *11*(1), 8-13. https://doi.org/10.1177/2043886920947110

Pallavicini, F. (2020). *Psicologia della realtà virtuale: Aspetti tecnologici, teorie e applicazioni per il benessere mentale*. Mondadori Università.

Pantano, E., & Servidio, R. (2012). Modeling innovative points of sales through virtual and immersive technologies. *Journal of Retailing and Consumer Services*, *19*(3), 279-286. https://doi.org/10.1016/j.jretconser.2012.02.002

Papagiannidis, S., Pantano, E., See-To, E. W., & Bourlakis, M. (2013). Modelling the determinants of a simulated experience in a virtual retail store and users' product purchasing intentions. *Journal of Marketing Management*, *29*(13-14), 1462-1492. http://dx.doi.org/10.1080/0267257X.2013.821150

Parnami, S., & Jain, T. (2021). A Study on Increase in the Usage of OTT Streaming Services. *International Journal of Research in Engineering, Science and Management*, *4*(8), 142-145.

Paun, R., & Olsen, D. (2022). Video content consumption during the pandemic: developing an innovative SVOD aggregator and communal viewing experience for the UK market. https://doi.org/10.25401/cardiffmet.19782712.v1

Peng, Y., & Ke, D. (2015). Consumer trust in 3D virtual worlds and its impact on real world purchase intention. *Nankai Business Review International*. http://dx.doi.org/10.1108/NBRI-03-2015-0009

Pham, T., Vermeulen, J., Tang, A., & MacDonald Vermeulen, L. (2018, June). Scale impacts elicited gestures for manipulating holograms: Implications for AR gesture design. In *Proceedings of the 2018 Designing Interactive Systems Conference* (pp. 227-240). https://doi.org/10.1145/3196709.3196719

Piumsomboon, T., Clark, A., Billinghurst, M., & Cockburn, A. (2013). User-defined gestures for augmented reality. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*(pp. 955-960). https://doi.org/10.1145/2468356.2468527

Plank, T., Jetter, H. C., Rädle, R., Klokmose, C. N., Luger, T., & Reiterer, H. (2017, May). Is Two Enough? ! Studying Benefits, Barriers, and Biases of Multi-Tablet Use for Collaborative Visualization. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 4548-4560). https://doi.org/10.1145/3025453.3025537

Porteous, J., Cavazza, M., and Charles, F. (2010). Applying planning to interactive storytelling: Narrative control using state constraints. *ACM Transactions on Intelligent*, *1*(2):111–130. http://dx.doi.org/10.1145/1869397.1869399

Poushneh, A., & Vasquez-Parraga, A. Z. (2017). Discernible impact of augmented reality on retail customer's experience, satisfaction and willingness to buy. *Journal of Retailing and Consumer Services*, *34*, 229-234. https://doi.org/10.1016/0022-4359(94)90013-2

Preece, J., Sharp, H., & Rogers, Y. (2004). *Interaction design. Oltre l'interazione uomo-macchina*.

Pynta, P., Seixas, S. A., Nield, G. E., Hier, J., Millward, E., & Silberstein, R. B. (2014). The Power of Social Television: Can Social Media Build Viewer Engagement?: A New Approach to Brain Imaging of Viewer Immersion. *Journal of Advertising Research*, *54*(1), 71-80. https://doi.org/10.2501/JAR-54-1-071-080

R Core Team. (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from https://www.R-project.org/

Regenbrecht, H., & Schubert, T. (2021). Measuring presence in augmented reality environments: design and a first test of a questionnaire. *arXiv preprint arXiv:2103.02831*. https://doi.org/10.48550/arXiv.2103.02831

Rogers, R. W. (1975). A protection motivation theory of fear appeals and attitude change. *The Journal of Psychology*, 91, 93– 114. https://doi.org/10.1080/00223980.1975.9915803

Romano, B., Sands, S., & Pallant, J. I. (2021). Augmented reality and the customer journey: an exploratory study. *Australasian Marketing Journal*, *29*(4), 354-363. https://doi.org/10.1016/j.ausmj.2020.06.010

Ronchi, E., Kinateder, M., Müller, M., Jost, M., Nehfischer, M., Pauli, P., et al. (2015). Evacuation travel paths in virtual reality experiments for tunnel safety analysis. *Fire Saf. J.* 71, 257–267. https://doi.org/10.1016/j.firesaf.2014.11.005

Roth, C. (2015). Experiencing interactive storytelling. *Vrije University, Amsterdam.*

Roth, C., & Koenitz, H. (2016, October). Evaluating the user experience of interactive digital narrative. In *Proceedings of the 1st International Workshop on Multimedia Alternate Realities* (pp. 31-36). https://doi.org/10.1145/2983298.2983302

Roth, C., & Koenitz, H. (2019, June). Bandersnatch, yea or nay? Reception and user experience of an interactive digital narrative video. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for T.V. and Online Video* (pp. 247-254). http://dx.doi.org/10.1145/3317697.3325124

Rovira, 2009 Rovira, A. (2009). The use of virtual reality in the study of people's responses to violent incidents. *Frontiers in Behavioral Neuroscience*, *3*(59), 1–10. https://doi.org/10.3389/neuro.08.059.2009

Rovira, A. (2009). The use of virtual reality in the study of people's responses to violent incidents. *Front. Behav. Neurosci.* 3:59. https://doi.org/10.3389/neuro.08.059.2009

Ruggieri, S., Biondi, A., & Boca, S. (2011). Il funzionamento psicosociale dei gruppi che interagiscono via web. *PSICOLOGIA SOCIALE*, *6*(1), 21-49. https://doi.org/10.1482/34472

Russell, J. A., & Pratt, G. (1980). A description of the affective quality attributed to environments. *Journal of personality and social psychology*, *38*(2), 311. https://doi.org/10.1037/0022-3514.38.2.311

Ryan, R.M., Rigby, C.S. and Przybylski, A. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion*. 30, 4 (2006), 344–360. https://doi.org/10.1007/s11031-006-9051-8

Sacks, R., Perlman, A., and Barak, R. (2013). Construction safety training using immersive virtual reality. *Construct. Manage. Econ.* 31, 1005–1017. https://doi.org/10.1080/01446193.2013.828844

Sastre-Merino, S., Nuñez, J. L. M., Pablo-Lerchundi, I., & Nufiez-del-Rio, C. (2020, December). Training STEAM educators in the COVID-19 emergency situation: *Redesigning teaching. In 2020 Sixth International Conference on e-Learning* (econf) (pp. 72-75). IEEE. https://doi.org/10.1109/econf51404.2020.9385461

Schatz, R., & Egger, S. (2008, March). Social interaction features for mobile TV services. In *2008 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting* (pp. 1-6). IEEE. http://dx.doi.org/10.1109%2FISBMSB.2008.4536629

Scholz, J., & Smith, A. N. (2016). Augmented reality: Designing immersive experiences that maximize consumer engagement. *Business Horizons*, *59*(2), 149-161.

https://doi.org/10.1016/j.bushor.2015.10.003

Schrader, C., & Bastiaens, T. J. (2012). Educational computer games and learning: The relationship between design, cognitive load, emotions and outcomes. *Journal of Interactive Learning Environments*, 23(3), 251–271.

Schrepp, M. (2015). User experience questionnaire handbook. *All you need to know to apply the UEQ successfully in your project*. http://dx.doi.org/10.13140/RG.2.1.2815.0245

Sermet, Y., and Demir, I. (2019). “Flood action VR: a virtual reality framework for disaster awareness and emergency response training,” in *ACM SIGGRAPH 2019 Posters, SIGGRAPH* 2019, (Los Angeles, CA). https://doi.org/10.1145/3306214.3338550

Shamma, D. A., Bastea-Forte, M., Joubert, N., & Liu, Y. (2008). Enhancing online personal connections through the synchronized sharing of online video. In *CHI'08 extended abstracts on Human factors in computing systems* (pp. 2931-2936). https://doi.org/10.1145/1358628.1358786

Shendarkar, A., Vasudevan, K., Lee, S., and Son, Y. J. (2008). Crowd simulation for emergency response using BDI agents based on immersive virtual reality. *Simul. Model. Pract. Theory* 16, 1415–1429. https://doi.org/10.1016/j.simpat.2008.07.004

Sila, I. (2013). Factors affecting the adoption of B2B e-commerce technologies. *Electronic commerce research*, *13*, 199-236. https://doi.org/10.1007/s10660-013-9110-7

Skinner, C. (2020). Flash Flood!: a SeriousGeoGames activity combining science festivals, video games, and virtual reality with research data for communicating flood risk and geomorphology. *Geosci. Commun.* 3, 1–17. https://doi.org/10.5194/gc-3-1-2020

Slater, M. (2003). A note on presence terminology. *Presence connect*, 3(3), 1-5.

Slater, M., Linakis, V., Usoh, M., & Kooper, R. (1996, July). Immersion, presence and performance in virtual environments: An experiment with tri-dimensional chess. In *Proceedings of the ACM symposium on virtual reality software and technology* (pp. 163-172). https://doi.org/10.1145/3304181.3304216

Smith, S., & Ericson, E. (2009). Using immersive game-based virtual reality to teach fire-safety skills to children. *Virtual reality*, *13*(2), 87-99. http://dx.doi.org/10.1007/s10055-009-0113-6

Spagnolli, A., Guardigli, E., Orso, V., Varotto, A., & Gamberini, L. (2014). Measuring user acceptance of wearable symbiotic devices: validation study across application scenarios. In *Symbiotic Interaction: Third International Workshop, Symbiotic 2014, Helsinki, Finland, October 30-31, 2014, Proceedings 3* (pp. 87-98). Springer International Publishing. https://doi.org/10.1007/978-3-319-13500-7_7

Spittle, B., Frutos-Pascual, M., Creed, C., & Williams, I. (2022). A review of interaction techniques for immersive environments. *IEEE Transactions on Visualization and Computer Graphics*. https://doi.org/10.1109/tvcg.2022.3174805

Statista (2021). *eCommerce - Europe*, Statista, retrieved April 19, 2022, from www.statista.com/outlook/dmo/ecommerce/europe.

Statista (2021)S. Lock s., “Online restaurant delivery growth worldwide 2019-2020, by country. https://www.statista.com/statistics/1238955/digital-restaurant-food-delivery-growth-in-selected-countries-worldwide/ (accessed May 10, 2022).

Statista,(2020). Worldwide retail e-commerce sales of Amazon from 2017 to 2021.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*, New York: Springer Science & Business Media.

Tabbaa, L., Ang, C. S., Siriaraya, P., She, W. J., and Prigerson, H. G. (2020). A reflection on virtual reality design for psychological, cognitive and behavioral interventions: design needs, opportunities and challenges. *Int. J. Hum. Comput. Interact.* 37, 851–866. https://doi.org/10.1080/10447318.2020.1848161

Tanis, M., & Postmes, T. (2007). Two faces of anonymity: Paradoxical effects of cues to identity in CMC. *Computers in Human Behavior*, 23, 955–970. http://dx.doi.org/10.1016/j.chb.2005.08.004

Tarnanas, I., and Manos, G. C. (2001). Using virtual reality to teach special populations how to cope in crisis: the case of a virtual earthquake. *Stud. Health Technol. Inform.*, 81, 495–501. https://doi.org/10.3233/978-1-60750-925-7-495

Tarre, K., Williams, A. S., Borges, L., Rishe, N. D., Barreto, A. B., & Ortega, F. R. (2018, November). Towards first person gamer modeling and the problem with game classification in user studies. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (pp. 1-2). https://doi.org/10.1145/3281505.3281590

Teng, M. Q., Gordon, E., and Hodge, J. (2019). “Participatory design of a virtual reality-based reentry training with a women's prison,” in *Conference on Human Factors in Computing Systems – Proceedings* (Glasgow), 1–8. https://doi.org/10.1145/3290607.3299050

Terraciano, A., McCrae, R. R., & Costa Jr, P. T. (2003). Factorial and construct validity of the Italian Positive and Negative Affect Schedule (PANAS). *European journal of psychological assessment*, 19(2), 131. https://doi.org/10.1027//1015-5759.19.2.131

Toraldo, M. L., Islam, G., & Mangia, G. (2018). Modes of knowing: Video research and the problem of elusive knowledges. *Organizational Research Methods*, 21(2), 438-465. https://doi.org/10.1177/1094428116657394

Van Elzakker, C. P., Delikostidis, I., & van Oosterom, P. J. (2008). Field-based usability evaluation methodology for mobile geo-applications. *The Cartographic Journal*, 45(2), 139-149. https://doi.org/10.1179/174327708X305139

Van Ginkel, S., Ruiz, D., Mononen, A., Karaman, C., de Keijzer, A., and Sitthiworachart, J. (2020). The impact of computer-mediated immediate feedback on developing oral presentation skills: an exploratory study in virtual reality. *J. Comput. Assist. Learn.* 36, 412–422. https://doi.org/10.1111/jcal.12424

Van Ittersum, K., & Feinberg, F. M. (2010). Cumulative timed intent: A new predictive tool for technology adoption. *Journal of Marketing Research*, *47*(5), 808-822.

Viola, E., Solari, F., & Chessa, M. (2022). Small Objects Manipulation in Immersive Virtual Reality. In *VISIGRAPP (2: HUCAPP)* (pp. 233-240). https://doi.org/10.5220/0010905200003124

Wang, M., Ye, Z. M., Shi, J. C., & Yang, Y. L. (2021, March). Scene-context-aware indoor object selection and movement in VR. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*(pp. 235-244). IEEE Computer Society.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54(6), 1063. https://doi.org/10.1037/0022-3514.54.6.1063

Wedel, M., Bigné, E., & Zhang, J. (2020). Virtual and augmented reality: Advancing research in consumer marketing. *International Journal of Research in Marketing*, *37*(3), 443-465. https://doi.org/10.1016/j.ijresmar.2020.04.004

Weisz, J. D., & Kiesler, S. (2008, October). How text and audio chat change the online video experience. In *Proceedings of the 1st international conference on designing interactive user experiences for T.V. and video* (pp. 9-18). https://doi.org/10.1145/1453805.1453809

Weisz, J. D., Kiesler, S., Zhang, H., Ren, Y., Kraut, R. E., & Konstan, J. A. (2007, April). Watching together: integrating text chat with video. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 877-886). https://doi.org/10.1145/1240624.1240756

WHO. (n.d.). *Floods*. https://www.who.int/health-topics/floods#tab=tab_1

Williams-Bell, F. M., Kapralos, B., Hogue, A., Murphy, B. M., and Weckman, E. J. (2015). Using serious games and virtual simulation for training in the fire service: a review. *Fire Technol.* 51, 553–584. https://doi.org/10.1007/s10694-014-0398-1

Wisniewski, Rebekka; Dennik-Champion, Gina; Peltier, James W. (2004). Emergency Preparedness Competencies. *JONA: The Journal of Nursing Administration*, 34(10), 475–480. https://doi.org/10.1097/00005110-200410000-00009

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence,* 7(3), 225-240. https://doi.org/10.1162/105474698565686

Witmer, B. G., Jerome, C. J., & Singer, M. J. (2005). The factor structure of the presence questionnaire. *Presence: Teleoperators & Virtual Environments*, 14(3), 298-312. https://doi.org/10.1162/105474605323384654

Wodehouse, A., & Abba, M. (2016). 3D visualisation for online retail: factors in consumer behaviour. *International Journal of Market Research*, *58*(3), 451-472. https://doi.org/10.2501/IJMR-2016-027

Wohlin, C., Höst, M., & Henningsson, K. (2006). Empirical research methods in Web and software Engineering. In *Web engineering* (pp. 409-430). Springer, Berlin, Heidelberg.

Wrzesien, M., Alcañiz, M., Botella, C., Burkhardt, J. M., Bretón-López, J., Ortega, M., et al. (2013). The therapeutic lamp: treating small-animal phobias. IEEE Comput. Graph. Appl. 33, 80–86. https://doi.org/10.1109/MCG.2013.12

Wymer, S. A., & Regan, E. A. (2005). Factors influencing e-commerce adoption and use by small and medium businesses. *Electronic markets*, *15*(4), 438-453. http://dx.doi.org/10.1080/10196780500303151

Xu, Q., & Sundar, S. S. (2014). Lights, camera, music, interaction! Interactive persuasion in e-commerce. *Communication Research*, *41*(2), 282-308. https://doi.org/10.1177/0093650212439062

Zaalberg, R., & Midden, C. J. H. (2013). Living behind dikes: Mimicking flooding experiences. *Risk Analysis*, *33*(5), 866–876. https://doi.org/10.1111/j.1539-6924.2012.01868.x